MAY 5 1930

JOURNAL
of the
Optical Society of America

VOLUME 20
No. 5

MAY
1930

Editor-in-Chief
PAUL D. FOOTE
Gulf Research Laboratory
327 Craft Avenue, Pittsburgh, Pa.

Assistant Editor-in-Chief
F. K. RICHTMYER
Cornell University,
Ithaca, N.Y.

CONTENTS

The Development of Thomas Young's Theory of Color Vision – – – – – – – – – – – – – – – – SELIG HECHT 231

On the Refraction of a Liquid Between Two Thin Prisms – – – – – – – – – – MORTON MASIUS AND W. E. LAWTON 271

The Amount of Ultraviolet Radiation Necessary to Cure Rickets – – – – – – – – GEORGE H. MAUGHAN AND J. A. DYE 279

Modes of Vibration of a Round Plate Cut from a Quartz Crystal – – – – – – – – – – – – – – – A. M. SKELLETT

Graphic Symbols in Applied Optics – – – – NOEL DEISCH 303

Index to Advertisers – – – – – – – – – – Cover III

ASSOCIATE EDITORS

Term Ending Dec. 31, 1931	*Term Ending Dec. 31, 1933*	*Term Ending Dec. 31, 1935*
FRANK ALLEN	SAUL DUSHMAN	H. E. IVES
K. T. COMPTON	K. S. GIBSON	R. A. MILLIKAN
K. K. DARROW	L. O. GRONDAHL	A. H. PFUND
G. S. FULCHER	P. E. KLOPSTEG	C. SHEARD
I. C. GARDNER	W. B. RAYTON	L. W. TAYLOR
L. A. JONES	H. B. WILLIAMS	H. S. UHLER

Published monthly by the OPTICAL SOCIETY OF AMERICA
Office of Publication
GEORGE BANTA PUBLISHING COMPANY, 450 Ahnaip St., Menasha, Wis.

JOURNAL of the Optical Society of America

Volume 20 **Number 5**

Published monthly by the Optical Society of America. Devoted to original papers on experimental, theoretical, and applied optics. Manuscripts for publication are invited, whether authors are members of the Optical Society or not.

Annual subscription $4.00 dating from January of current year; or $5.00 in combination with THE REVIEW OF SCIENTIFIC INSTRUMENTS. Single copies $0.60.

BACK NUMBERS

THE JOURNAL OF THE OPTICAL SOCIETY OF AMERICA
Volume 1-5, (1917 to 1921) inclusive, per set.....$16.00

THE JOURNAL OF THE OPTICAL SOCIETY OF AMERICA AND REVIEW OF SCIENTIFIC INSTRUMENTS
Volume 6 and 7, (1922 and 1923), per volume.....$ 5.00
Volume 8 to 19, two volumes per year, (1924 to 1929), inclusive, per volume.....$ 2.50

THE JOURNAL OF THE OPTICAL SOCIETY OF AMERICA
Volume 20— (1930-), one volume per year, per volume.....$ 4.00

Address publishers, "George Banta Publishing Company, 450-454 Ahnaip Street, Menasha, Wisconsin," or "F. K. Richtmyer, Managing Editor, Cornell University, Ithaca, New York," for all correspondence regarding subscriptions, back numbers, changes of address, advertising rates and accounts, and missing numbers.

THE OPTICAL SOCIETY OF AMERICA

President—L. A. Jones
Eastman Kodak Company
Rochester, N.Y.

Vice-President—E. C. Crittenden
Bureau of Standards
Washington, D.C.

Secretary—L. B. Tuckerman
Bureau of Standards
Washington, D.C.

Treasurer—Adolph Lomb
Rochester, N.Y.

The Council consists of the above officers, the Editor-in-Chief and the Assistant Editor-in-Chief of the JOURNAL, the past President, I. G. Priest, Bureau of Standards, Washington, D.C., and the following elected members: A. H. Compton, University of Chicago, Chicago, Ill.; I. C. Gardner, Bureau of Standards, Washington, D.C.; P. E. Klopsteg, Central Scientific Company, Chicago, Ill.; E. D. Tillyer, American Optical Company, Southbridge, Mass.

The object of the Optical Society of America is to serve the interests of all who are engaged in any branch of optics from fundamental research to the manufacture of optical goods.

The Constitution provides that any one who has contributed materially to the advancement of optics shall be eligible to regular membership in the Society with the privilege of voting and holding office. Any individual or corporation interested in optics is eligible to associate membership. Annual dues including subscription to the JOURNAL are five dollars for individual associate membership, seven dollars and fifty cents for individual regular membership and fifty dollars for corporation membership. Further information may be obtained from the secretary.

Rochester, N.Y., Section of the Optical Society of America

President—T. R. Wilkins
Vice-President—Otto Sandvik
Secretary—Donald L. Wood
Treasurer—E. M. Lowry

Councilors: F. M. Bishop, H. A. Clark, F. C. Fairbanks, H. F. Kurtz

---

Entered as second-class matter May 31, 1921, at the post office at Menasha, Wisconsin, under the Act of March 3, 1879. Acceptance for mailing at special rate of postage provided for in the Act of February 28, 1925, embodied in paragraph 4, section 412, P. L. and R., authorized June 19, 1925.

Journal
of the
Optical Society of America

Vol. 20 MAY, 1930 No. 5

THE DEVELOPMENT OF THOMAS YOUNG'S THEORY OF COLOR VISION*

By Selig Hecht

I. THOMAS YOUNG'S FUNDAMENTAL CONTRIBUTION

It has been known for a good deal over a hundred years that the sensation produced by nearly any given light may be reproduced by the mixture in proper proportions of the light from three selected portions of the spectrum. More generally, and perhaps more accurately stated, all color sensations produced by light may be described in terms of three properly chosen variables. One may say

$$L = pA + qB + rC \tag{1}$$

in which L is the light, A, B, and C are the chosen variables, usually spoken of as primaries, and p, q, and r, are the coefficients representing the respective amounts of the primaries required to match the sensation produced by the light. The particular specifications of these primaries may vary, and indeed may be quite arbitrarily assumed. Nevertheless, they are adequate as a formal description of the properties of color vision insofar as they relate to color mixing.

Such a basic fact should obviously be utilized and incorporated into any mechanism proposed as a theoretical basis for color vision. This indeed was first done about 130 years ago by Thomas Young, the centennary of whose death we are today commemorating. Young[43] supposed that there are three kinds of fibers in the retina, each produc-

* The greater part of this paper was given as an address at the Thomas Young Memorial Meeting of the Optical Society of America, held on October 24th, 1929, at Cornell University, Ithaca, to commemorate the centennary of the death of Thomas Young.

ing a characteristic sensation, one of red, another of green, and a third of violet (or blue). Each type of fiber is sensitive practically to the whole visible spectrum but the first possesses a maximum of sensibility in the red, the second in the green, and the third in the blue. Various color sensations then result from the relative strengths with which the three different fibers are stimulated by the objective light.

Newton[33] had supposed that certain elements in the retina vibrated in unison with the incident light, and that the "harmony and discord of colors arise from the proportion of the vibrations propagated through the fibers of the optic nerve into the brain as the harmony and discord of sound arise from the proportions of the vibrations of the air."

This conception was not acceptable to Thomas Young. In his famous Bakerian lecture to the Royal Society in 1801 he therefore suggested that since "it is almost impossible to conceive each sensitive point on the retina to contain an infinite number of particles each capable of vibrating in perfect unison with every possible undulation, it becomes necessary to suppose the number limited, for instance, to the three principal colors, red, yellow, and blue, of which the undulations are related in magnitude nearly as the numbers 8, 7, and 6; and that each of the particles is capable of being put in motion less or more forcibly by undulations differing less or more from a perfect unison; for instance, the undulations of green light being nearly in the ratio of 6 1/2 will affect equally the particles in unison with yellow and blue and produce the same effect as a light composed of those two species; and each sensitive filament of the nerve may consist of three portions, one for each principal color." In a later communication, due to the erroneous observations of Wollaston, Young[31] suggested that these basic color sensations are red, green, and violet.

Young does not specifically say that the sensation of white is produced by the simultaneous stimulation of the three fibers. However, it is apparent that this is an obvious conclusion from his theory. Koenig[32] has pointed out that Thomas Young rests his theory on a principle which we in physiology at the present time recognize but which at that time had not yet been enunciated by its author, Johannes Müller.[32] This is the principle of the specific energies of nerves which supposes that no matter how a given nerve fiber is stimulated, it always produces a certain "energy" or quality which is associated with it. In this case a specific fiber will produce the sensation of red, or of green, or of blue depending upon its own self only and not upon the particular wave length which is used in stimulating it.

It is not usually recognized that Thomas Young's insight into the problem of color vision went even deeper than this. He suggested that the color confusions which were made by his contemporary, Dalton, were due to the lack or paralysis of those fibers in the retina which were concerned with the red sensation.

For nearly fifty years these ideas lay buried. They were rescued from oblivion by Maxwell[30] and by Helmholtz[14] almost at the same time. It would be difficult to find a better statement of the position in the history of science of this theory and of Thomas Young's work in general than the one given by Helmholtz in his Popular Scientific Lectures.[15] Helmholtz after describing various aspects of color vision says: "The nature of colors with all these marvelous and complicated relations was a riddle which Goethe in vain attempted to solve; nor were we physicists and physiologists more successful. I include myself in the number for I long worked at the task without getting any nearer my object until I discovered that a wonderfully simple solution had been made at the beginning of this century and had been in print ever since for anyone to read who chose. This solution was found and published by the same Thomas Young who first showed the correct method of interpreting the Egyptian hieroglyphics. He was one of the most acute men who ever lived, but he had the misfortune to be born far in advance of his contemporaries. They looked on him with astonishment but they could not follow his bold speculations. And so a mass of his most important thoughts remained buried and forgotten in the Philosophical Transactions of the Royal Society until a later generation slowly arrived at the rediscovery of his discoveries and came to appreciate the force of his argument and the accuracy of his conclusions."

II. THE ADEQUACY OF YOUNG'S IDEA

Few among us today suppose that Young's idea as here given in its simple form or as elaborated by Helmholtz,[14] Koenig,[22] von Kries,[27] and others, is adequate as a *complete* theory for the mechanism of color vision. We know very definitely that it is not. The question, however, has often been raised as to whether it can serve even as the basis for a theory of color vision. This question was raised by Aubert,[3] and by Hering[17] almost immediately after Maxwell and Helmholtz first brought Young's idea to light. Aubert, Hering, and more recently Ladd-Franklin[28] considered the problem of color theory from a totally different point of view than did Young. Instead of relying on the objective facts of color mixtures which state that three variables are enough to

describe color sensation, these investigators started from their own sensations of color. Perhaps the simplest way of indicating their point of view is to consider the sensations of yellow and white. Mixtures of green and blue lights give a continuous series of colors, the blue-greens, in which both colors are identifiable. The same is true with mixtures of red and blue lights; these give a series of violets and purples in which one can see the two components. But in mixtures of red and green, a new sensation, yellow, arises which contains neither red nor green. Yellow seems to be something by itself. It is definitely not red plus green; and we recognize this by saying that a color is greenish yellow or reddish yellow, not greenish red nor reddish green. The same is true for mixtures of yellow and blue or of green, red, and blue which give white, a totally new sensation.

If Young's idea is correct then it must be supposed that yellow is a central phenomenon occurring in the brain when in the retina the red and green receiving fibers function simultaneously; similarly, white is the sensation which occurs in the brain when all three receiving fibers, the red, the green, and the blue, function in the retina. It is precisely against such a formulation that Aubert, Hering, and Ladd-Franklin ranged themselves. And as a result first Aubert, then Hering, and later, Ladd-Franklin devised theories which take into consideration the curious uniqueness of the yellow and the white sensations.

One must not minimize the difference between these two positions. It would perhaps be difficult to do so because the work of the last fifty years in color vision has had its origins mainly in the antagonism which has existed between those who, like Helmholtz, Koenig, von Kries, Abney, and others, have accepted and amplified Young's original formulation and those who, like Hering, Aubert, and Ladd-Franklin, have held opposing views. Nevertheless, it is important to remember that these two sets of views are not necessarily antagonistic when, as suggested by v. Kries, they are considered as partial descriptions of different parts of the process of color vision as a whole. It is only when they are concerned with a formulation of the *receptor* process in the retina that the two views are in sharp contrast. And here their differences demand an experimental test.

If there is to be developed an adequate theory for the mechanism of color vision a decision must be made in the very beginning as to which of the two conceptions of yellow and white is correct. Are there special substances or processes in the retina for the reception of yellow and white or are yellow and white phenomena which arise in the brain out

of the impulses coming from the three kinds of fibers or processes in the retina? Is a three receptor notion adequate or do we need to have a four or perhaps a five receptor system to account for color vision. Since the data of color mixing may be adequately described on a three dimensional basis they can be described on a four dimensional basis as well,—a fact which has been frequently pointed out.[5,13,24,37] The problem is, therefore, essentially a physiological one, and not a mathematical one.

If red light and green light fall on a given retinal area of one eye and a yellow sensation results, it is not possible to decide whether this is the result of a stimulation of two receptors or of one receptor. But if red light falls on the retina of one eye, and green light falls on the corresponding portion of the retina of the other eye, and the result is a yellow sensation, then only Young's idea is tenable, because there must be two receptors involved in making the yellow sensation from red and green. This is precisely what happens in the binocular mixing of colors.

The binocular mixing of colors is an old story. Curiously enough, Helmholtz was unable to fuse colors binocularly and the weight of his authority prevented the binocular fusion of colors from a proper acceptance. Nevertheless, Hering and a number of other investigators showed precisely what was wrong with Helmholtz' experiments and arranged apparatus to overcome it. Largely as a result of the work of Trendelenburg[40] and recently of Rochat,[26] there is no doubt any longer that the binocular fusion of colors takes place and that one can get from this method regular color mixing equations resembling those secured monocularly. I have recently[12] described a very simple method of demonstrating binocular fusion of colors which renders it impossible for anyone to doubt its existence and which could be adopted as a routine laboratory experiment.

It must be obvious that if one can produce a white sensation by putting yellow in one eye and blue in the other, there is no use in talking of a special receptor for white. White and yellow are sensations which are produced in the brain out of the impulses that come in from the three receptors in the retina. We may, therefore, conclude that the uniqueness of yellow and white as sensations is no obstacle to adopting Thomas Young's three receptor idea as a basis on which to build a theory for the receptor mechanism of color vision. This does not mean that the ideas which prompted Aubert, Hering, and the others may not be applicable to some other portion of the complicated series of events

involved in color vision. That, however, is beside our present point. For the receptor process a three fiber idea is adequate.

III. THE NATURE OF YOUNG'S THREE FIBERS

Having adopted the notion of three receptor elements as a basis for color vision, we are obliged to consider the further specification of the nature of these three receptors. It has frequently been suggested, and with excellent reason, that the three fibers may be considered as three species of cones. In this way there is united the fiber idea of Young with the substance idea of Helmholtz. Each type of cone contains its own photosensitive substance, and is connected with an optic nerve fiber, stimulation of which produces in the brain the corresponding sensation. Let us adopt this notion as a clarifying and a simplifying one, and proceed to the necessary next step in the theoretical treatment of color vision, namely, the quantitative description of the spectral properties of these three receiving elements.

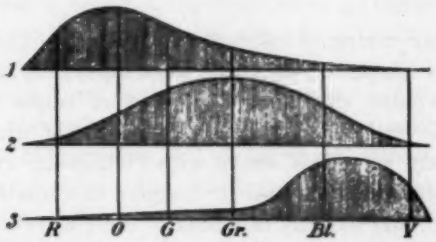


FIG. 1. *Spectral sensibility distribution of Young's three primaries as suggested by Helmholtz (14). The shapes of the curves are entirely diagrammatic.*

Efforts to define them have not been lacking. Young originally gave no precise description of his three fibers except to say that their maxima were respectively in the blue, the green, and the red, and that they each were sensitive to the whole spectrum. However, Helmholtz[14] made a drawing of their spectral properties, which, though entirely fanciful, expressed Young's and his notions of them. This drawing is reproduced in Fig. 1. Later, the work of Maxwell,[30] of Koenig and Dieterici,[26] of Abney,[1] and recently of Wright[42] has given us three curves currently known as excitation curves or sensation curves which are supposed to represent the spectral characteristics of the three primaries at the basis of color mixing. These curves are reproduced in Fig. 2. They give the values derived by Weaver from the work of Koenig and Dieterici and of Abney, and have been adopted by the

Optical Society[41] as standard primaries. The areas under the three curves are made equal following Koenig's original suggestion that they contribute equally in making white. In terms of them, there is usually constructed the familiar color triangle.

Do these three excitation curves really represent the physiological primaries of Thomas Young? Are they a description of the spectral properties of three fundamental processes concerned with the reception of color? There is a tacit assumption among us that they are, though most of us would never acknowledge this publicly. Certainly Koenig who first derived them accurately felt that they were very nearly the descriptions of the three sensations though he knew they were not exactly so. Therefore, by making a few linear and homogeneous trans-

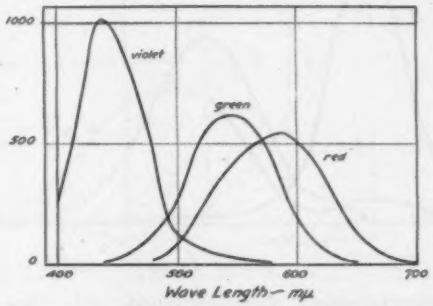


FIG. 2. *The excitation curves adopted as standard by the Optical Society of America. The curves are taken from the Report of the Colorimetry Committee (41) and were derived by Weaver by recomputing and averaging the data of Koenig and Dieterici (26) and Abney (1). They correspond essentially to what Koenig and Dieterici called their Elementarempfindungen.*

formations in their values, he derived from them three curves, which he called the basic sensation curves, or Grundempfindungen, and which he thought might represent these primaries physiologically. These are given in Fig. 3. Koenig never forgot that his three Grundempfindungen were transformed color mixture equations; and when it became necessary later to change their form he did not hesitate to do so.[24] Since Koenig's time the original Grundempfindungen have acquired an air of great respectability; so much so, that the very necessary and important changes, which he made in them after their first formulation, have been consistently ignored and are never referred to in the papers or books of the last thirty years. Koenig always remembered that the Grundempfindungen, as well as the excitation curves from which they are derived, represent nothing more than the data of color

mixture. As such they were brilliant derivations because they showed that the color mixture equations given by the three main kinds of color blind eyes could each be described by means of only two of the three Grundempfindungen which the normal eye yields.

It is quite true that color mixture is one of the most important characteristics,—perhaps the most important characteristic of color vision. Nevertheless, it is not the only significant property of color vision. If the excitation curves or the Grundempfindungen are really a description of the spectral characters of the three physiological primaries of Thomas Young, they should furnish an adequate theoretical basis for

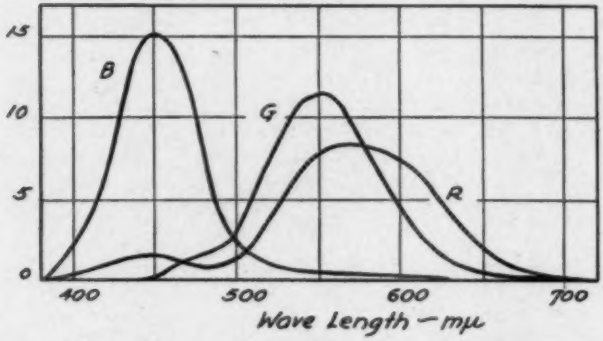


FIG. 3. *The Grundempfindungen of Koenig and Dieterici (26). These curves were derived by Koenig and Dieterici by combining in various ways their Elementarempfindungen as given in Fig. 2.*

the other properties of color vision as well. Unfortunately, the Grundempfindungen have not served in this capacity. It is usually not recognized that the Grundempfindungen were shown to be inadequate in this respect almost immediately after their formulation. This was done by Helmholtz[18] in his treatment of Koenig's data of hue discrimination. Other failures have since become apparent.

I believe that we can learn much from the closer examination of these failures of the excitation curves. Their inadequacies may perhaps suggest some other specifications for the three physiological primaries of Thomas Young in terms of which one may describe not only the data of color mixtures, but of other properties of color vision as well.

IV. THE BRIGHTNESS OF THE SPECTRUM

It is apparent that the only straightforward way of stating the data of color mixture is in terms of the actual energies of the three

primaries required to match a light of given wave length and known energy content. To my knowledge this very desirable method of presenting the data has never been used. The nearest approach to such a formulation is to refer the data to a known, standard spectrum, and to express the color of different portions of that spectrum in terms of the three primaries taken from that same spectrum. With some modifications, this method of presentation has recently been used for the first time by Wright.[42] Fig. 4 reproduces Wright's data, in which the rela-

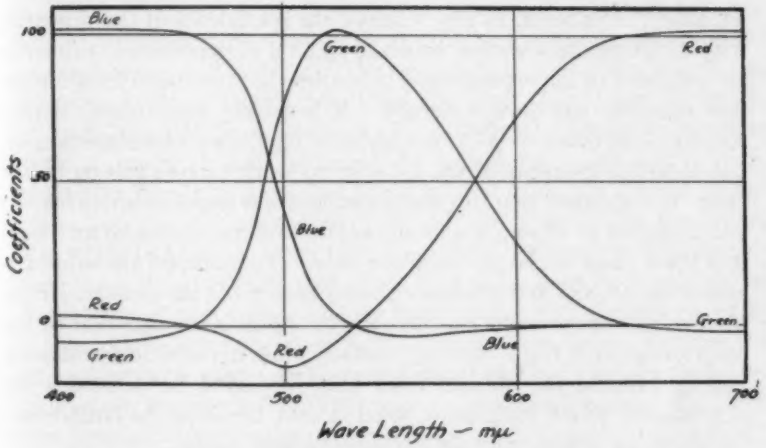


FIG. 4. *The direct data of color mixture. This figure is made from the date of Wright (42) and represents the findings of 10 observers.*

tions of the primaries are shown with simplicity and directness. The values of each of the primaries in this spectrum have been placed arbitrarily at unity; but since their relative brightnesses are given, it is possible to compute the energy values throughout and to make the necessary conversions to an energy basis.

It is a familiar fact that a given light, whether mixed or physically monochromatic, produces not merely a sensation of color but a sensation of brightness as well. It is the particular virtue of the energy method of giving the data of color mixture, as exemplified in the modified form used by Wright, that the presentation is independent of luminosity. In contrast to this is the presentation of the same material in terms of the excitation curves or of the Grundempfindungen, because these take their form in addition from the luminosity of the spectrum. If, instead

of giving merely the relative amounts of the primaries required to match a given part of a standard spectrum, there are given the actual amounts, it is apparent that the curves given in Fig. 4, for example, would become modified to include the luminosity of the different parts of the spectrum relative to the assumed values of the primaries. Such curves do not record the actual luminosity of the spectrum because the luminosities of the primaries have been arbitrarily kept at unity.

We may assume as Koenig did, that the actual luminosity values of the primaries cannot be determined. He, therefore, kept them entirely arbitrary. And since, as Fig. 4 shows, the coefficients of the primaries may sometimes be negative, he made up a set of hypothetical primaries so composed of the experimental primaries that the coefficients for the new primaries are always positive. It is usually not recognized that Koenig made this step in the formation of his Elementarempfindungen; but a careful following of the steps in his work[25] shows this to be the case. It is apparent that this linear homogeneous transformation may be accomplished by so simple a means as lowering the horizontal axis until it is lower than the largest negative value. To maintain the arbitrariness of the old and new primaries Koenig then made the areas under the three curves equal to one another, and the result is the three excitation curves as given in Fig. 2. Koenig tried still other hypothetical primaries, and by keeping the data of color blinds in mind he derived a set of primaries, which when made equal in area, he called the Grundempfindungen, to which reference has already been made.

The reason that decided Koenig to make the areas under the excitation curves equal in area was more than the arbitrary one of not knowing what they really are. He made them equal because he thought they contributed equally in the making of white so far as their color qualities were concerned. He knew that a yellow and a blue which combined to make white were very different in brightness; but since he thought they were equal in color power, he made them equal. In this way he dodged the really fundamental problem, which we, however, must face. Koenig avoided it by assuming the primaries to be equal in color power while knowing them to be unequal in brightness. Such a paradox gains nothing by being accepted as an assumption; rather it must be considered as vigorously demanding an explanation.

The difficulty which the three excitation curves or their derived Grundempfindungen present us with is as follows. They include the data of color mixture; and in terms of them color equations may be derived. One may thus get out of the data what one puts into them. But even

though they include luminosity data, they do so only in a partial way, and no information can be derived from them which helps us in understanding or deriving luminosity equations without the help of additional data. In other words the excitation curves, though distorted by luminosity considerations, do not determine luminosity.

If the three curves, which are equal in area, are added together they certainly do not give a brightness distribution of the spectrum which at all resembles the correct brightness distribution of the spectrum. The blue curve is obviously large out of all proportion. In order to achieve a correct summation of the curves it is necessary to make luminosity measurements of the three primaries, which demands the use of heterochromic photometry. The first effort at such an evaluation was made by Brodhun[4] for his color blind eye. But Abney[1] was the first to evaluate the relative magnitude of the three excitations required to make a proper visibility curve of the normal eye. This was later repeated by Ives,[18] by Exner[8] and others. And from this work there has come out the not unexpected fact that these areas are not equal. According to Ives,[18] in order to make a white curve whose total brightness is equal to 1.000 the area of the blue curve must be multiplied by 0.006, that of the green curve by 0.426 and that of the red curve by 0.568. Exner's values are slightly different from these. Perhaps the most acceptable values are those computed by Judd[20] for the excitation curves adopted by the Optical Society and the visibility curve of Gibson and Tyndall. Judd's values are 0.015 violet, 0.547 green, and 0.438 red.

These are all unimpeachable facts. But they make it extremely difficult to formulate a physiological picture of what they mean, especially in the matter of treating the excitation curves or the Grundempfindungen as the real physiological primaries which must accomplish what Thomas Young's notion expects them to. If areas are any indication of the function or the capacity for stimulation of these three primaries then blue is insignificant. This is indeed fact because yellow needs only an insignificant amount of blue in order to be made white. The blue is exceedingly effective in changing the color but it adds nothing to the brightness. If, for example, one assumes three kinds of cones whose sensibility characteristics correspond to the three sensation curves then equal areas would mean an equal number of three kinds of cones because brightness is determined among other things by the number of retinal elements functioning.[12a] But if the proper luminosity coefficients are applied then it would seem that there are almost no

blue cons eand that the green and red cones are roughly in the ratio of 3 to 4.

Let us consider the meaning of this somewhat more closely. There are two ways of treating the luminosity coefficients of the excitation curves or Grundempfindungen in their relation to primary receptors. The first is to assume that the three kinds of cones which they represent are present in the retina in an approximate ratio of 1:1:1. According to Ives' luminosity coefficients one must, therefore, suppose that to make white the ratio of the cones stimulated is 1 blue to 75 green to 100 red. If this were so, then as white light is made more intense it should change in hue. It should become greener and finally bluer because there are relatively more green cones and more blue cones available which would be brought into play by the higher intensities. This change in hue of course is not true. On the other hand, suppose that anatomically the actual ratio of the three kinds of cones is 1:75:100. Then it should never be possible to make a blue light as bright as a bright red one because there never would be enough blue cones to match the number of red ones. However, we know that the only limit to the brightness of a blue light is the amount of energy that can be secured from the source of illumination.

My point is simply that the three excitation curves, because they are derived from color mixture data can describe color mixture. The partial or complete introduction into them of luminosity data only confuses matters, and as Ives has already shown, decreases their practical usefulness. It certainly makes it difficult to think of them as representing the three receptors first thought of by Young; and it becomes difficult to think of them in terms of any physiological mechanisms with which we are already acquainted.

V. THE VISUAL ACUITY OF THE COLOR BLIND

The inadequacy of the excitation curves as a description of the three necessary receptors for a color theory first became evident to me in trying to understand certain results which we secured in measuring the visual acuity of color blind people. Our reasons for undertaking the measurements were somewhat as follows. It has been shown that the visual acuity of the eye is proportional to the number of cones which are functionally present in a unit area of the retina.[12a] The tri-receptor idea assumes that there are three kinds of cones. Suppose for the moment that they occur in the ratio of 1:1:1. Thomas Young's explanation of color blindness may then be considered to mean that a red-blind or a

green-blind person lacks one set of cones, the red or the green. Therefore a dichromat, since he possesses only two-thirds the normal quota of cones, should have a maximum visual acuity of only two-thirds normal. Moreover, because of the spectral distribution of the excitation curves, the visual acuity of a dichromat should be low in the spectral region which corresponds to the lost set of cones, and should be normal in the rest.

Our measurements, however, showed no such condition. In the first place, the red-blind and green-blind individuals whom we tested showed in white light a maximal visual acuity of the same magnitude as normal eyes. This by itself is not very serious. One may adopt an old idea of Fick,[10] which was accepted to a certain extent both by Helmholtz and by Koenig, that a color blind does not really lose the function of one set of cones. What happens rather is that the sensitive substance in the affected cones assumes the spectral characteristics of one of the other kinds of cones. If, for example, in a red-blind person the red-receiving cones should have their substance take on the spectral properites of the substance in the green-receiving cones then such a person would still have the full quota of cones. But there would be a confusion of red and green, both of which would appear as yellow. This is because these lights would stimulate the red-receiving cones and the green-receiving cones to exactly the same extent. The impulses going up from these cones would, however, still register red and green, and since the impulses would be simultaneous and equal the brain would synthesize yellow. Such a person would be color blind, but since he has the normal number of cones per unit retinal area, his visual acuity in white light would be of the same magnitude as that of the normal trichromat. Indeed, the fact that we found dichromats to possess a normal visual acuity in white light lends considerable support to Fick's representation of the basis of color-blindness.

Our next findings, however, were much more disturbing, because while they further supported Fick's idea, they were not understandable in terms of the notion that the excitation curves represent the three kinds of cones. Consideration of Fig. 2 shows that if the red curve, for example, dropped out, then the visual acuity should be much reduced in the red end of the spectrum, less so in the green part, and hardly at all in the blue region. We found, however, that whereas the visual acuity of a red-blind is really lower than normal in the red region of the spectrum, it is distinctly higher than normal in the other regions. Obviously the red-receiving cones are not lost; their spectral sensibility

is merely changed to correspond with that of the green-receiving ones. But closer consideration of the data, as well as of the three excitation curves presents difficulties at once. Doubling the green curve at the expense of the red curve will result in a loss of the red, an increase in the green, and hardly any change in the blue. Our measurements, however, showed only a slight rise above normal in the green, but a decided increase in the blue corresponding in size to about the loss in the red.

If we try to explain this by supposing the red-receiving substance to take on the properties of the blue-receiving substance we are met with the fact that this does not correspond to the sensations of the color blind as given to us by monocular color-blind individuals. Moreover, we accentuate the paradoxical situation with which we had to deal in the last section with regard to the brightness of the spectrum. Suppose, for example, that the blue, green, and red cones are present anatomically in the retina in the ratio of 1:1:1. Then as we saw before we would need 1 blue to 75 green to 100 red cones to make white. If now in the case of a red blind person the red cones assume the spectral sensibilities of the blue cones then white light should call out 1 real blue cone, 75 green cones, and 1 red cone with the blue receiving properties. And white light would then appear bluish-green to a red blind. We know this is not true from cases of people who are color blind in one eye only. Moreover, in such a case the total number of cones stimulated would be in the ratio of 77 for the color blind compared to 176 for a normal; and the visual acuity of a dichromat in white light should be less than half of normal. This we know not to be true; it is just like normal.

If we take the other supposition that the cones are present anatomically in the ratio of 1 blue to 75 green to 100 red, we are no better off. In the case of our dichromat, what would happen to the 100 red cones whose receiving substance has been changed from red to blue? There would be a large drop in visual acuity in the red as expected; but there would be a relatively enormous increase in the visual acuity in the blue. The matter may be put more effectively if we consider a case with just a trivial loss of red, say 10 per cent. The drop in visual acuity in the red would be insignificant; but the increase in the blue should be fantastically large. However, our experiments show that the decrease in visual acuity in the red is just about compensated by the increase in the blue.

If the excitation curves really described the properties of the physiological primaries which we have here spoken of as cones, then these

various findings should be consistent and sensible. As it is they are as confusing as many other phenomena of color vision have been, and point to the fact that the excitation curves are not the correct description of anything else except color mixture data.

VI. THE SATURATION OF THE SPECTRUM

Brightness and visual acuity are not the only properties of color vision which raise difficulties in the interpretation of the excitation curves in terms of physiological mechanism. There are many others. It will be enough to present two more, because they are typical. Let us consider the problem of saturation.

It has been known for a long time[14] that the relative brightness of complementary colors differ greatly depending upon the complementary pairs. For example, to make white with the two lights of wave length 609 $m\mu$ and 494 $m\mu$ they must be very nearly the same brightness; for the two complementaries 576 $m\mu$ and 475 $m\mu$ to make white the yellow must be about ten times as bright as the blue-green; and finally for the yellow and the extreme violet to make white they must be in the ratio of very nearly 100:1.

These facts are well known and have been recently again studied adequately by Sinden.[35] Their meaning, however, is linked to a very interesting speculation which I believe is due to Helmholtz.[14] It is familiar that homogeneous spectral colors are not equally saturated and that some of them produce more white sensation than others. Helmholtz supposed that the reason that yellow and violet require such a disproportionate brightness is that yellow is much less saturated than violet. In the violet most of the sensation produced is violet, whereas in the yellow only a tiny fraction of the sensation is yellow and the rest is white. In order for these complementaries to make white one must take enough of the yellow so that the color portion of it will be the same magnitude as the color portion of the violet; and this obviously means a low brightness of violet and a high brightness of yellow. There is no question that this suggestion of Helmholtz makes the data of complementary colors conform with our common experience. One speaks of spectral red and of spectral blue as saturated colors, whereas yellow-green, yellow, and orange are considered unsaturated colors.

Recently Priest and Brichwedde[35] have investigated this problem of the saturation of the spectrum from a more direct point of view. They took a photometric field of white light of a given brightness. To one side they added homogeneous spectral light while taking out as much white

as necessary in order to keep the brightness of the two sides constant. They then determined the minimum amount of spectral color that had to be added in order that one half of the field just appeared colored. They found that there is necessary a great deal more of yellow, of orange, and of yellow-green to produce this minimal color perception than of red, of blue, or of violet.

The argument though more direct, is precisely the same as in the previous one for complementary colors. If a color is very saturated then a certain amount of it will be required to be added to the white light in order to produce a color sensation. But if a color is unsaturated and already carries with it a great deal of white producing capacity, then considerably more of it will be necessary to produce a minimal color sensation when it is added to white. Priest and Brickwedde's experiments show about the same distribution of saturation in the spectrum as the complementary color data of Sinden do.

According to the tri-receptor idea, white sensation means an equal stimulation of the red, the green, and the blue receptors. If the excitation curves describe the spectral properties of the three receptors, then the white producing capacity of any portion of the spectrum should become apparent from the spectral distribution of the three excitation curves. There should be a great deal of over-lapping of these three curves near the yellow and very little over-lapping in the red and blue. However, it needs only one glance at the three sensation curves in their equal area state to realize at once that the minimum saturation should not be in the yellow but some place in the blue-green. Moreover, in the three curves as corrected for brightness there is hardly any white at all since the blue curve has practically disappeared. We see, therefore, again that the three sensation curves do not help us very much in a physiological and realistic interpretation of Thomas Young's tri-receptor idea.

VII. HUE DISCRIMINATION AND THE SENSATION CURVES

Color vision may be defined as the capacity of the eye to divide the visible spectrum into a series of regions which produce qualitatively different sensory effects. The normal human eye can separate the visible spectrum with complete certainty into about 180 patches or hues which cannot be made to look like one another by varying their intensity. Measurements of hue discrimination are nearly a century old; but their modern aspects began with Koenig and Dieterici[25] and culminated in the work of Steindler,[39] of Jones,[19] and of Laurens and Hamilton.[29] This work has shown that the minimum distance in wave length

which it is necessary to move along the spectrum before one can tell that the hue is different varies in a striking but consistent way. There are two maxima of hue discrimination. One is at about 490 mμ and another at about 575 mμ. In addition there are two minor maxima of hue discrimination, one at about 430 mμ and the other at about 640 mμ. The precise shape of these curves varies surprisingly from person to person but their general form is the same. Fig. 5 shows the data of Steindler for her own eye.

It is apparent that the characteristics of hue discrimination should be an expression of the properties of the three receptors at the basis of the mechanism of color vision. Here again, Helmholtz[16] was the first, and I believe the only one, to suggest a quantitative relation between

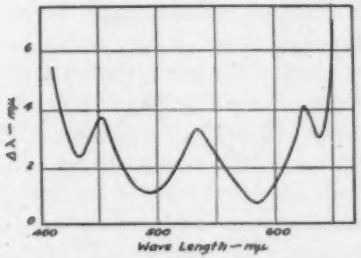


FIG. 5. *Hue discrimination in the spectrum. The curve is taken from Steindler (39), and represents the data for her own eye.*

the two. He supposed that the hue change in the spectrum is determined by the relative rates of change which the spectral characteristics of the three receptors undergo in the spectrum. The precise mathematical formulation proposed by Helmholtz is somewhat insecure especially where it depends on the Weber-Fechner law which he himself had shown to be inadequate. But the basic notion is undoubtedly sound.

It is, therefore, an interesting commentary on the general usefulness of the Grundempfindungen that when Helmholtz tried to compute the hue discrimination data of Koenig and Dieterici in this way, he had to discard their Grundempfindungen and invent a set which bore no resemblance to them. For historical interest these are reproduced in Fig. 6. Even if they do reproduce Koenig and Dieterici's hue discrimination data, it is apparent that the curves are not very useful otherwise. The fact that each curve possesses two maxima in the spectrum is difficult to accept easily. The additional fact that they cross each other at such odd places is hard to reconcile with other knowledge. Never-

theless, these curves possess one interesting property: they assume the spectrum to be much less saturated than could be supposed from the ordinary excitation curves.

Since Helmholtz' time all attempts to derive hue discrimination data from either the excitation curves or from the Grundempfindungen have not been much more successful (cf. Judd[31]). I may say that I spent a good deal of time doing the same thing myself with no success, before I realized that others had also failed. Helmholtz' idea is undoubtedly correct fundamentally. The conclusion to be drawn from its failure is,

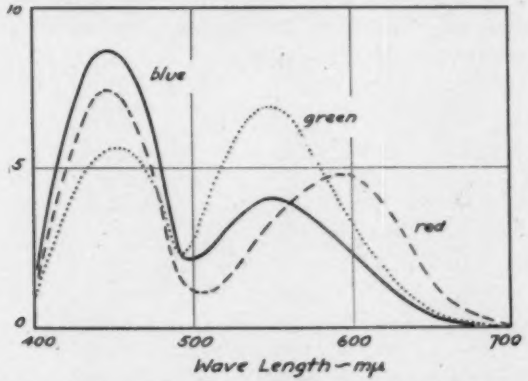


FIG. 6. *The three fundamental primaries which Helmholtz (16) had to assume in order to describe the data of Koenig and Dieterici (25) on hue discrimination. These primaries are formed by combinations of the excitation curves or Elementarempfindungen such as are given in Fig. 2. The curves are copied in part from Peddie (34).*

I believe, the same which has been emphasized in the previous sections of this paper. It is that, though the excitation curves and the Grundempfindungen describe the data of color mixture, they do not describe the characteristics of the three necessary physiological primaries postulated by Thomas Young.

VIII. REQUISITES FOR CHARACTERIZING THE PRIMARIES

What I have said up to now will be readily agreed to by most people who have worked in color vision and have tried to think realistically of its physiological implications. It is not that we do not possess significant data. It is true that we need more varied data with the eye of a single investigator in order to give a really final quantitative statement of the relations among different aspects of vision. But most of the data of color vision are good enough; some are better than are to be found in

other branches of physics and physiology. What we need is some sort of unified hypothesis which should integrate the information which we possess, and serve as an incentive for more accurate and varied work. It is not enough for the data of color mixture to describe quantitatively the data of color mixture, which is what the excitation curves accomplish. What is necessary is a quantitative description of the three receptors postulated by Thomas Young, in terms of which it should be possible not merely to derive the data of color mixture, but the data of brightness, the data of saturation, the data of hue discrimination, and the data of any other relevant aspect of color vision. It is with the *quantitative* aspects of these phenomena that we should be concerned. The day is long past when anyone should be permitted to propound literary theories of vision.

It may not be amiss to enumerate the various requisites for an adequate theoretical description of the receptor mechanism of color vision. In general such a description will be realistic and physiological. This probably means that it will consider the three receptors as three kinds of cones each possessing a specific photosensitive substance and each connected with a specific nerve. The cones will behave like other sense cells and will conform to what is known of cells in general. The nerves will act like other nerves, which means that they will have what Johannes Müller called "specific energies"; they will probably behave according to the all-or-nothing law; and their properties will correspond to the knowledge of nerves accumulated by physiologists such as Adrian,[2] Forbes,[11] Erlanger and Gasser[7] and many others. Specifically, an adequate theory will include all the color mixing equations. It will straighten out the brightness difficulty and the visual acuity muddle; it will solve the saturation problem; and will explain the hue discrimination in the spectrum. It will remember the neutral points of color blinds. It will even bear in mind the special sensations associated with different parts of the spectrum by the normal and color-blind eye.

I recognize that this is a large order and that it will require the labor of many minds to formulate it. At the same time, I wish to call attention to it on this occasion, because it is the most important defect in the present-day study of color vision. A science without a working theory is a disorganized mass without a guiding spirit. I have tried to show that Thomas Young's tri-receptor idea can serve as the nucleus for such a working hypothesis. Unfortunately the efforts of later workers to make this idea quantitative have led only to blind alleys, because they have relied too much on the data of color mixture alone. It has been neces-

sary to clear away some of these misapprehensions before the problem could be faced. Is it possible perhaps to suggest some developments of Young's idea which are more adequate?

In the last four years I have examined various ideas and suggestions, most of which have had to be discarded. Some portions of them, however, have remained. And in our laboratory, for purposes of thinking and discussion, we have tentatively adopted certain quantitative notions as a basis of guidance for our future work. Though I have discussed these ideas freely with most people who have been interested in the matter, I should never have presented them publicly under ordinary circumstances. Today, however, is not an ordinary occasion. It is the extraordinary occasion on which we have gathered to honor the memory not merely of a great man of Science, but of the man who by his brilliant insight laid the basis for precisely the thing we need most at the present time, namely, an adequate theory for the receptor process concerned in color vision. And so if I have accepted the invitation of your Committee to present these ideas to you, it is not because I believe these ideas to be adequate. As you shall see, I know that they are not. It is with considerable hesitation that I present them to you; but I do so only if you will consider them as an experiment in quantitative thinking. If I were a musician I should perhaps have composed a set of "Variations on a Theme by Thomas Young" for this day. Take this theme and variations as a contribution to this occasion.

IX. A POSSIBLE HYPOTHESIS

Let us suppose that there are present in the retina three kinds of cones; blue-receiving cones, green-receiving cones, and red-receiving cones. Let us further suppose that in the fovea they are present in approximately equal numbers. The spectral properties of these cones will be described presently. At first it is well to understand what we mean by the first assumption. We suppose that a given cone, possessing a photosensitive substance whose absorption is normally greater in the blue, or in the green, or in the red part of the spectrum is joined to a nerve fiber which is so connected with the brain that whenever the photosensitive substance in the cone is changed by light and starts an impulse in the nerve, that nerve will register respectively blue or green or red in the brain. It means moreover that no matter what method may be used to start this impulse, no matter what the wave length of the light may be, indeed no matter what the photosensitive substance in the cone may be, an impulse travelling up a "blue" nerve will

register blue in the brain, a "red" nerve red, and a "green" nerve green.

The spectral characteristics of these three types of cones,—practically their absorption spectra,—are given in the curves in Fig. 7 and in detail in Table 1. It is apparent that these three curves are radically different from both the excitation curves of Fig. 2 and the Grundempfindungen of Koenig in Fig. 3. Since it is their quantitative properties

TABLE 1. *Primaries V, G, and R tentatively proposed as a description of the three receptors postulated by Young as a basis for color vision. V, G, and R add up directly to give Ives' luminosity curve L.*

λ	*V*	*G*	*R*	*L*
410	2.5	1.4	0.1	4
420	5.9	3.7	0.5	10
430	9.9	6.1	2.0	18
440	15.3	9.3	4.4	29
450	21.5	12.8	7.7	42
460	28.1	17.6	12.4	58
470	38.7	29.5	21.8	90
480	53.5	48.8	35.7	138
490	75.0	78.5	58.6	212
500	113.6	124.6	97.8	336
510	144.6	175.4	160.1	480
520	196.7	229.9	211.4	638
530	248.9	284.3	262.9	796
540	289.2	325.1	303.7	918
550	313.9	348.4	329.8	992
560	321.6	346.2	331.2	999
570	310.2	325.6	317.2	953
580	278.1	299.4	301.5	879
590	234.0	256.3	267.7	758
600	190.7	212.7	229.6	633
610	146.9	161.4	181.7	490
620	103.0	119.4	139.7	362
630	69.4	81.7	99.0	250
640	44.4	52.4	67.2	164
650	26.5	30.3	43.2	100
660	15.8	16.6	27.6	60
670	9.9	10.0	19.1	39
680	4.8	5.0	12.2	22

which will be used to describe in detail various phenomena of color vision it is important to tell something about the origin and formulation of these new primaries. Otherwise the matter will appear too arbitrary to be useful for further work and thought.

Essentially they were specially designed to account for precisely those properties of color vision which we have found to be impossible of

description in terms of the excitation curves. Naturally, therefore, they had to include those properties of the excitation curves which are useful in the description of color mixture, and to a certain extent of color blindness. Like the excitation curves, our primary curves have certain specific crossing points. The red and the green curves cross at about 575 mμ to account for yellow; the violet and the green curves cross at about 490 mμ; the violet and the red curves cross at about 500 mμ. These take care of certain complementary color equations, as well as of the neutral points of the three types of color blind individuals. Like the excitation curves they are equal in area; but unlike them, they add up very simply to give the visibility curve of the spectrum of white light.

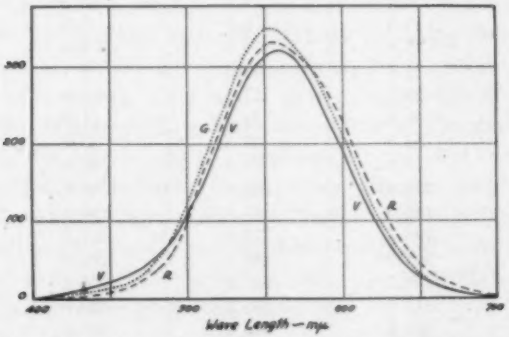


FIG. 7. *Spectral distribution of the primaries, V, G, and R, here tentatively proposed as a basis for a quantitative treatment of color vision.*

A study of the different properties of color vision shows that we may arrange them in a sort of hierarchy of magnitudes, each ascending magnitude being more restrictive than the other. If we accept the crossing points as stated above and demand that the areas of the curves be equal we can draw an infinite series of curves which will satisfy this demand. The excitation curves are just one of this infinite variety. The primary curves in Fig. 7 are another. If we demand in addition that the three equal primary curves shall add up to give a normal visibility curve we at once limit the positions of these curves, though there are still a large number of curves which will satisfy this added requirement. If we further restrict the shapes of these curves with the demand that they yield the data of color mixture as adequately as the excitation curves do, it is apparent that the possible types of curves, though still large, are more reduced than before. From now on any additional de-

mands that may be made of the primaries have a surprisingly restricting effect. For example, when the requirements of visual acuity are added the curves become so restricted that one may draw them free-hand to a first degree of approximation. Once these curves have been secured in this approximate way quantitative investigations may be begun on them to see how the various properties of color vision are derived from them, and how systematic changes in them yield different quantitative relations. I do not wish to minimize the labor involved in such a procedure; it is pretty large. Nevertheless it is worth recording that once the relations between the different sets of data of color vision were mastered and understood, the first approximate primary curves were not difficult to draw, and were not very dissimilar from those given in Fig. 7.

It is interesting to point out some peculiarities of these relations. For example, color mixture equations are not very different for different people. This is apparent from the extensive data of the past, but more particularly from the recent investigations of Wright shown in Fig. 4. On the other hand hue discrimination is a very delicate and distinctly individual affair. Steindler's data show this well, as do the data of Laurens and Hamilton. It is important to remember such things when one tries to invent a set of primaries to account for the data of color vision, because unfortunately, one has to use data from different individuals for the different functions. Individual differences will thus not be very apparent in color mixture equations but will show up strikingly in hue discrimination. Therefore in making the primaries it is well to satisfy the latter requirements first.

Before proceeding to show how the three primaries drawn in Fig. 7 describe the various data of color vision, I wish to call attention to one minor point. You will notice that although I have previously spoken of blue cones, the corresponding curve which I have presented in Fig. 7 is marked violet rather than blue. This is because violet seems to act as a unit in equations of color mixing and in several other things as well. For the moment, however, I wish to leave this paradox aside. I shall take it up later and show how it may be resolved.

X. COLOR MIXTURE AND BRIGHTNESS

The three primaries, *V*, *G*, and *R* which I have given in Fig. 7 and Table 1 must, first and foremost, account for the data of color mixture. The simplest way of showing this is to consider them in terms of the excitation curves of Fig. 2. These latter are the data of color mixture.

Therefore if it can be shown that the functions V, G, and R are linear derivatives of the functions v, g, r, described by the excitation curves, then any color equation involving v, g, and r, can just as well be written in terms of V, G, and R.

The functions V, G, and R refer to white light. The functions v, g, and r which now refer to an equal energy spectrum, must therefore be transformed so as to refer to white light. Calling v_0, g_0, and r_0 these transformed values of the excitation curves, then our primaries V, G, and R are related to them according to the following equations:

$$V=0.97(.030v_0+.570g_0+.400r_0) \quad (2)$$

$$G=1.03(.015v_0+.561g_0+.424r_0) \quad (3)$$

$$R=1.02(.006v_0+.426g_0+.567r_0) \quad (4)$$

The computations* indicated in these equations involve the following steps. The v, g, and r values are taken from Troland.[41] They are multiplied by the energy distribution (E) of Abbot-Priest sunlight[34a] in which the energy at 590 mμ is placed equal to 1.00. Then in order to make v_0, g_0, and r_0 equal in area and also equal in area to V, G, and R, we write $v_0=6.22\ Ev$; $g_0=6.23\ Eg$; and $r_0=6.15\ Er$. The values of v_0, g_0, and r_0 so secured are now introduced into equations (2), (3) and (4).

The values of V, G, and R calculated in this manner from the color mixture data as given by the excitation curves, v, g, and r, are shown in the points in Fig. 8. The smooth lines are the theoretical values V, G, and R, of Fig. 7 and Table 1. As is apparent to anyone who has used the excitation curves for similar calculations, the agreement between the theoretical curves and the experimental data is quite good, the differences being within a very few percent. Therefore the values of V, G, and R of Table 1 and Fig. 7 are an adequate description of the data of color mixture.**

* In these computations I have had the help of Mr. Simon Shlaer to whom I wish here to present my thanks.

** In preparing the present paper for the press, I have found two minor errors in the computations. The first is that I used the visibility curve for an equal energy spectrum instead of for white light. The difference between the two is slight and is well within the range of individual variation of the visibility curve for white light. The second error is that though I planned to make V, G, and R equal in area, and of course equal to v_0, g_0, and r_0, I find that this is not quite true. Calling the areas of v_0, g_0, and r_0 equal to 1.00, the area of V is 0.97, that of G is 1.03, and that of R is 1.02. Hence the extra coefficients in equations (2), (3), and (4). I have purposely avoided the rather tedious recomputations which are required to correct these two unimportant errors, for the reason that the ideas with which this paper is concerned are unaffected by them. *I consider all the calculations and all the numbers here given as distinctly of a preliminary and exploratory nature.*

Having established this, we are no better off than we were with the excitation curves of Fig. 2. We must proceed to the next step of including as well the brightness of the spectrum. This, however, is simple

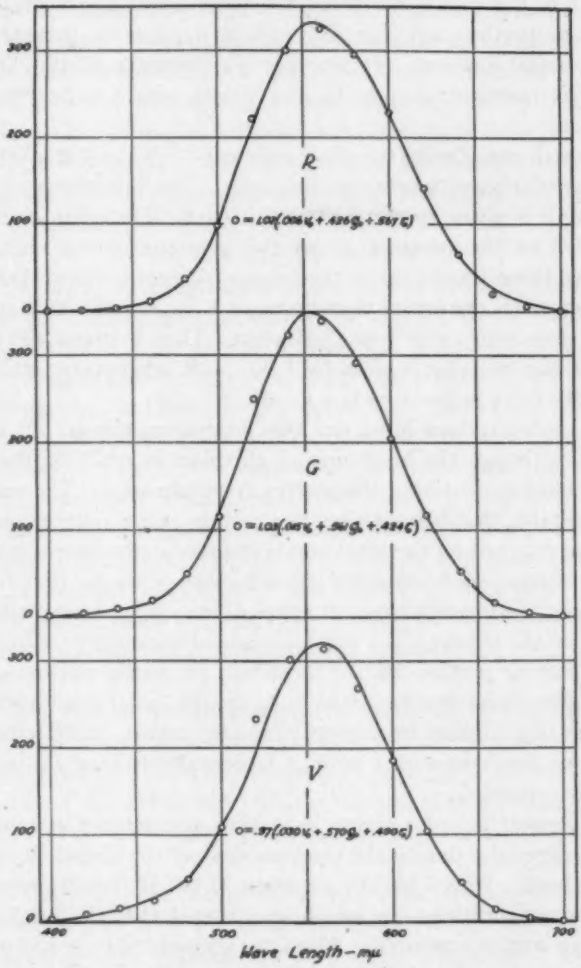


FIG. 8. *Relation of the primaries V, G, and R to the data of color mixture as included in the excitation curves of Fig. 2. The points are calculated from the excitation curves, v, g, and r, as shown; the curves are the primaries themselves. Not much weight is to be attached to the fact that the calculated point for 520 mμ is off the curve in all three primaries; this is entirely due to the visibility curve on the basis of which the three primaries are derived and which is peculiarly low at this particular spot.*

and straightforward. The three curves V, G, and R add up accurately and directly to make the normal visibility curve, because as Table 1 shows, they were constructed to do precisely this. Furthermore, since the areas of the curves are equal the three kinds of receptors whose spectral properties they describe can be considered present in the retina in equal numbers. In this way we eliminate all the difficulties which were raised in Sections IV and V with regard to brightness and visual acuity.

It is worth considering the three primaries V, G, and R a little more closely in relation to brightness and color. The brightness at a given wave length is given by the sum of $V+G+R$. The color, however, is determined by the amounts of any two primaries which remain over and above three times that of the lowest. This is because three times the lowest value represents those parts of V, G, and R which are equal and therefore result in a white sensation. Thus between 400 mμ and about 500 mμ the color is given by $V+G-2R$, whereas beyond that to 700 mμ the color is given by $G+R-2V$.

These considerations bring out two interesting things. In the first place, even though the brightness of all colors is additive, their color qualities tend to neutralize themselves in certain ways. The reason for this is probably that brightness is to a certain extent determined in the retina and depends on the total number of cones stimulated, regardless of color, whereas color sensation is much more of central nervous affair and depends on certain brain changes of which we know nothing at present. In the second place, a comparison of the relative values of the color producing portion and of the white producing portion of most wave lengths shows that the latter is by far the larger magnitude of the two in relation to total brightness. In other words, speaking approximately, the brightness of a color is largely determined by its white-producing capacity.

These aspects of color vision have been the subject of much controversy, especially during the vigorous days of the Helmholtz and the Hering schools. It was held by members of the Helmholtz school that the colors are additive,—by which they meant that the brightness of colors is an additive property. This is an undoubted fact, and is shown to make sense in terms of the behavior of our primaries. Apparently on the contrary, the Hering school maintained that colors are not additive, but tend to neutralize each other in certain ways, and that the brightness of a given color depends on the white-producing capacity associated with it—its white valence. The Hering school therefore deemed

it necessary to have separate processes,—at first these were separate receptor processes, later they were left unlocalized,—for color and for white or brightness. To physicists this notion of the neutralization of colors, and the association of brightness with whiteness has seemed a very fanciful and mystical thing. Nevertheless there has always been a significant point in these ideas. And it surprised me to see how they came out quite unexpectedly in a purely physical and physiological formulation of the primaries such as I have here presented. Brightness is additive; colors qualities do neutralize one another in certain ways; and the brightness of a color is very largely determined by its white-producing capacity.

XI. HUE DISCRIMINATION

We have described quantitatively the data of color mixture and of brightness in terms of the primaries V, G, and R. However, these aspects of color vision are obviously simple, and their solution represents only a very slight progress over what has been achieved with the usual excitation curves. As I have already pointed out, there are a large number of curves which are capable of doing precisely the same thing. What it is necessary to do now is to show that these three primaries describe adequately many additional data of color vision. I propose to include two more, hue discrimination and saturation, both of which we have already partly discussed. Though for purely didactic reasons they are presented here after color mixture and brightness, their treatment and derivation should really precede these because of their individual and therefore more restricting influences on the shapes of the three primaries V, G, and R.

The best data of hue discrimination which are now available are those of Laurens and Hamilton.[29] These were made at constant brightness. I have selected for study the data of Laurens' own eye, and since the primaries V, G, and R were constructed with them in mind, it should be possible to derive them from the primaries with a point for point agreement.

We have seen that the change in hue which the visible spectrum shows must be connected with the relative changes which the primary receptors show in their spectral characteristics. Helmholtz[16] assumed a non-linear and quite complicated relation between the two, and in order to derive the even meager quantitative data available in his day, he had to invent a set of primaries which were neither probable nor useful. I propose therefore to use a simple and direct formulation in terms of our primaries V, G, and R.

Let us suppose that the hue change will depend on the rates of change $dV/d\lambda$, $dG/d\lambda$, and $dR/d\lambda$ in relation to one another. Since V, G, and R always add up directly to control the brightness we need to consider at any point the relative rates of only the two most rapidly changing of the three, since the third is eliminated, as we saw in the last section, by making the brightness constant and equal. We therefore see that at constant brightness the change in hue, H, over the wave lengths between 400 mμ and 500 mμ will depend on the way the primaries V and G change in relation to each other; and that between 500$m\mu$ and 700$m\mu$ the change in hue will depend on the way the primaries G and R change in relation to each other. Thus from 400 to 500

$$\frac{dH}{d\lambda}=\frac{dV}{d\lambda}-\frac{dG}{d\lambda} \tag{5}$$

and from 500 to 700

$$\frac{dH}{d\lambda}=\frac{dR}{d\lambda}-\frac{dG}{d\lambda} \tag{6}$$

will describe the situation. These two equations may be simplified by writing them as

$$\frac{dH}{d\lambda}=\frac{d(V-G)}{d\lambda} \tag{7}$$

and

$$\frac{dH}{d\lambda}=\frac{d(R-G)}{d\lambda}. \tag{8}$$

To make the computations, we must remember first, that what is actually measured is not a differential $d\lambda$, but a finite difference $\Delta\lambda$; and second, that Laurens and Hamilton specifically measured a wave-length distance $\Delta\lambda_c$ necessary for a minimal difference in hue ΔH rather than the reverse. This difference in hue, ΔH, therefore represents a constant difference in the composition of the two perceptibly different hues in the amounts of V and G or of R and G which they contain at constant brightness. This may correspond to a difference in the number of cones of each species functional for the two hues. Writing $\Delta H = K$ the equations become

$$\Delta\lambda_c=\frac{K.\Delta\lambda}{\Delta(V-G)} \tag{9}$$

and

$$\Delta\lambda_c=\frac{K.\Delta\lambda}{\Delta(R-G)} \tag{10}$$

in terms of which the experimentally determined distance $\Delta\lambda_e$ along the spectrum required for a constant hue difference is proportional to the distance $\Delta\lambda$ which one has to move on the curves in Fig. 7 to yield a constant difference in V-G or in R-G.

I have made this computation from the curves in Figure 7 using a $\Delta\lambda$ of 10 $m\mu$ width. The values of $10/\Delta(V\text{-}G)$ or $10/\Delta(R\text{-}G)$ so obtained are referred to the center of this band, and when multiplied by $K=5$, reproduce the data for Laurens' eye with complete fidelity. Actually the computation can be made arithmetically in the simplest way. One writes down from Table 1 for every 10 $m\mu$ in the spectrum the value of V-G or of R-G. Then each value is subtracted from the next above. The reciprocals of the resulting differences are multiplied by 5, and the result

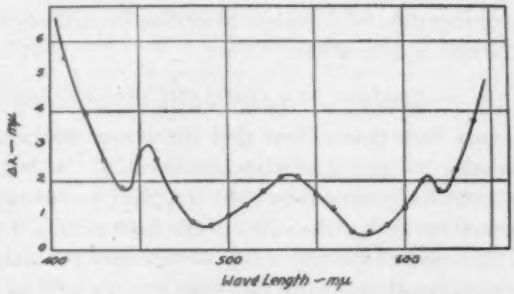


FIG. 9. *The data of hue discrimination of Laurens' eye taken from Laurens and Hamilton (29). The smooth curve gives the data, whereas the points are computed from the primaries V, G, and R.*

is the value which should correspond with $\Delta\lambda_e$ experimentally determined at λ midway between the two points 10 $m\mu$ apart which were used for the computation.

The way in which computed and experimental values agree is shown in Fig. 9, where the continuous curve represents the data for Laurens' eye, and the points are the computed ones. This agreement should not excite one too much, because as I have already said, the curves V, G, and R were constructed with a view toward making them yield these very data. What is significant about the agreement, however, is that the values of V, G, and R, would have to undergo only the most minor changes in order to fit any other set of hue discrimination data. These minor changes will affect only slightly the color-mixture equations, and not at all the areas of the curves, nor their summation to form a visibility curve.

The method which I have used to derive the hue discrimination data shows clearly the hierarchy of magnitudes of which I spoke in Section IX. It is important to dwell on this for a moment. The first order of magnitude in the description of V, G, and R is their brightness. This corresponds to the areas under the three receptor curves, and are made equal to begin with. The second order of magnitude is the shape of the three primary curves. The forms of the curves take care of the data of color mixture, and changes may be made in their shape to bring them into conformity with the data of color mixtures without affecting their areas. The third order of magnitude, in reality the second differential of the area, is concerned with hue discrimination. In this way the slightest changes in the curves though they affect hue discrimination will hardly affect the data of color mixture. This explains why hue discrimination may vary considerably from person to person even though their color mixture equations may be similar.

XII. SATURATION AND COMPLEMENTARY COLORS

It has already been pointed out that the curves which describe the three primaries, V, G, and R overlap considerably. It thus seems that much of the sensation produced by light of a given wave length is white, and only a small portion of the effect of the light results in a sensation of color. This seems surprising when we consider the richness of the color sensations associated with the spectrum as well as with other lights. It is necessary therefore to show that the data of saturation such as those secured by Priest and Brickwedde[35] demand just such a partition of color sensation and white sensation.

I will be pardoned if I refer to the sequence of events which actually took place in connection with this particular point. After the V, G, and R curves had been drawn in an approximate form, I felt it necessary to test this saturation idea before going any further, because I deemed the point crucial. I therefore started to work on the complementary color data of Sinden to see whether they could be understood in terms of the V, G, and R curves. It soon became apparent, however, that though complementary color data could undoubtedly be used as corroborative evidence as I shall presently show, they were not useful as crucial evidence. This is because there is a whole stretch in the middle of the spectrum between about 500 $m\mu$ and 600 $m\mu$, for which one cannot get monochromatic complementaries; and it is precisely in this middle stretch that very important changes in saturation take place. It was at this point that Priest and Brickwedde's experiments became available to me

through Dr. Priest's kindness. The data corresponded so surprisingly well with expectation from the rough V, G, R curves that I was sufficiently satisfied with the primaries to make no changes in them until the present occasion. For presenting them now I have made a few slight changes in the original, rough V, G, R curves, so that the ones given now in Fig. 7 yield Priest and Brickwedde's data with practically perfect fidelity.

It will be recalled that Priest and Brickwedde's experiments consist in determining the minimum amount of spectral light which must be added to one side of a photometric field of uniform white of a given brightness in order to produce a barely perceptible color sensation in the white. If we divide the brightness of the added spectral color by the total brightness we get a fraction which Priest and Brickwedde have called the least perceptible colorimetric purity.

It must be apparent that if light of a given wave length were completely saturated, a certain amount of it will have to be added to white in order for the white to appear colored. Furthermore, if light of a different wave length, which is only half saturated, is used, then twice as much of it would be needed in order for the same amount of "color" to be added as before. From the curves which I have proposed for the three receptors it is a simple matter to compute for each wave length its saturation. This is the ratio of the color part to the total brightness. Thus in the region between λ400 and λ500 the color portion is $V+G-2R$, and in the region between about λ500 and λ700 it is $G+R-2V$. The total brightness is $V+G+R$; therefore the saturation, S_λ, is the ratio of these two quantities, and

$$S_\lambda = \frac{V+G-2R}{V+G+R} \tag{11}$$

for λ400 to λ500, and

$$S_\lambda = \frac{G+R-2V}{V+G+R} \tag{12}$$

for λ500 to λ700. The values of $S\lambda$ can be computed directly from Fig. 7 and Table 1.

In Priest and Brickwedde's experiments let B_λ be the brightness of the homogeneous light which needs to be added to secure a just perceptible color; let B_w be the brightness of the constant white field, and let P_λ be the minimum perceptible colorimetric purity. Then

$$P_\lambda = \frac{B_\lambda}{B_w} \tag{13}$$

and represents the definition of colorimetric purity. Since the amount of "color" added is assumed constant, the quantity B_λ which has to be added to furnish this color is inversely proportional to the saturation S_λ. This may be written

$$B_\lambda = \frac{k}{S_\lambda} \tag{14}$$

and can be introduced into equation (13) which now becomes

$$P_\lambda = \frac{K}{S_\lambda} \tag{15}$$

since B_w is kept experimentally constant. In other words the minimum perceptible colorimetric purity as determined by Priest and Brickwedde is inversely proportional to the saturation as determinable from the values of the primaries V, G, and R.

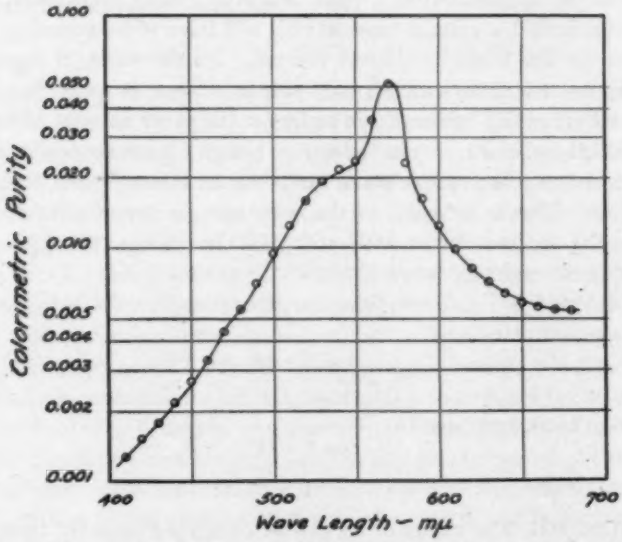


FIG. 10. *Saturation of the spectrum as found by Priest and Brickwedde (35) in terms of what they call the least perceptible colorimetric purity. The points are those computed from the primaries V, G, and R, whereas the continuous line represents the average data of Priest and Brickwedde. I am grateful to Dr. I. G. Priest for his gracious permission to use the data before their publication in final form.*

Using equation (15), (11) and (12), the values of V, G, and R from Fig. 7 and Table 1, and putting $K = 0.0012$, I have computed the values of P_λ. The results are given in Fig. 10. Here the continuous line repre-

sents the data of Priest and Brickwedde, while the circles are the values calculated from theory. It is apparent that the agreement between experiment and expectation is practically perfect. That is as it should be because the curves V, G, and R were made so that they would yield this agreement. However, as already pointed out, the changes required to make them so were insignificant.

Before concluding this section I wish to present one more relationship which has already been referred to, namely, complementary colors and their relative brightness. These computations I present just as they were made. It will be apparent that no changes need be introduced because of them.

The relation between complementary color pairs has been known for many years and has been frequently determined. Essentially such determinations are merely problems in color mixture and since our primaries V, G, and R have been shown to be adequate in describing color mixture data, their capacity to reproduce complementary pairs would not be anything additional. However, complementaries exhibit certain peculiarities, already referred to, with regard to their relative brightness in making white. These measurements relate to saturation. Therefore, an adequate computation of complementaries in terms of their relative brightness would furnish a corroborative test for the sufficiency of V, G, and R as a preliminary description of the three receptors required in the theory of color vision.

For this purpose we possess the most recent and also the most complete set of measurements by Sinden.[28] Sinden determined for 12 complementary pairs the relative brightness of each member of a pair required to make a white equal in hue and brightness to a standard white of about 75 photons. Sinden actually gives the ratios L_1/L_2 of the luminosities for each pair; but knowing that they must add up to 75 units it is simple to compute the separate brightness of each member of the pair. These values are given in Table 2. For each wave length in the Table we may secure its composition in terms of V, G, and R from Fig. 7. The sum of these, $V+G+R$, gives the relative luminosity of the wave length. Let us now multiply the values of V, G, and R for each color by such a factor, that their sum will arbitrarily equal the numerical value of the luminosity which Sinden found in his experiments and which are given in Columns 3 and 4 of Table 2. Thus V_1, G_1, and R_1 will be the values for one member of a pair, and V_2, G_2, and R_2, the values of the other member. Then, since in white $V=G=R$, the sum V_1+V_2 should

equal G_1+G_2, and equal R_1+R_2, each of which should have a value of 25.0 since their sum is equal to 75.

Table 2 gives the values for the twelve pairs computed in this way. Two things should be noticed. First it is clear that V_1+V_2, G_1+G_2 and R_1+R_2 are equal to one another to within 3 or 4 percent as an average. The worst value is the first V_1+V_2 which is 8 percent off. The experiments themselves are only good to about this extent. Hence the agreement is adequate. But second, the values are consistent for all the pairs. Thus the V_1+V_2 values among themselves, the G_1+G_2 values among themselves, and the R_1+R_2 values among themselves agree to a sur-

TABLE 2. *Composition of complementary colors in terms of the Primaries, V, G, and R.*

Complementaries		Luminosities		Composition		
λ_1	λ_2	L_1	L_2	V_1+V_2	G_1+G_2	R_1+R_2
650	496	31.6	43.4	23.2	25.8	26.0
609	493.5	39.3	35.7	24.3	26.3	24.3
591	490	50.3	24.7	24.3	26.2	24.7
586	487.5	55.4	19.6	24.3	26.0	24.7
580	482.5	62.0	13.0	24.5	25.8	24.7
578.5	480.5	64.2	10.8	24.5	25.8	24.7
576.5	477.5	65.8	9.2	24.7	25.6	24.7
575.5	474.5	67.7	7.3	24.7	25.6	24.7
574	472	69.0	6.0	24.7	25.6	24.7
573	466.5	70.6	4.4	24.7	25.4	24.7
572	459	72.1	2.9	24.9	25.4	24.7
570.5	443	73.2	1.8	24.7	25.6	24.7

prising extent. What this undoubtedly means is that the white which Sinden used is not the same white which I have used for making my values of V, G, and R. His is apparently about 4 or 5 percent greener; or rather to match his white it is necessary to have a ratio of about $1.00V$ to $1.04G$ to $1.00R$; this is also within the errors of heterochromic photometry. It must, therefore, be concluded that Sinden's data may be quantitatively accounted for on the basis of the V, G, and R curves which I have drawn as representing the spectral characteristics of the three receptors concerned in color vision.

XIII. COLOR BLINDNESS

Reference has already been made to color blindness, and before concluding this paper it is necessary to develop this aspect to a slight extent. It may as well be confessed that color-blindness is the subject that I

have thought least about, and therefore it is in the least satisfactory condition. Nevertheless it is necessary to discuss the matter in order to point out the weak points as well as the strong ones.

It has already been shown that we must extend the original notion of Thomas Young that color blindness means the simple loss or inactivity of one of the three kinds of receptors. This was apparent to Fick, as well as to Koenig and to others. What we know requires us to adopt the notion that in color blindness the spectral characteristics of one of the receptors is changed so that it becomes identical with one of the others. For example let us suppose that the photosensitive substance in the green-receiving cone becomes changed so that its absorption spectrum (and probably its chemical constitution) becomes identical with that in the red-receiving cones. Under such circumstances the changed green-receiving cone and the red-receiving cone will be equally stimulated by any light. However, the rest of the mechanism remains undisturbed; that is, the "green" fiber will still cause a green sensation in the brain as before. Therefore there will be no way for such an eye to discriminate hues in the spectrum between about 550 $m\mu$ and 700 $m\mu$, all this region appearing predominantly yellow. That this is true we know first from the reports of those people who are color blind in one eye and normal in the other; and second from the studies on the hue discrimination of color blinds by Steindler and by Laurens and Hamilton. Moreover, since the heights of the V and the R curves are identical at about 490 $m\mu$, the curve of the changed "green" substance, since it is like the "red" will also cross the V curve at the same spot. But, since the sensations produced will still be violet, green, and red in equal amounts the result will be a white sensation, or the neutral point in the spectrum of the green-blind.

The same situation arises if the substance in the "red" cones take on the absorption spectrum and the photochemical properties of the substance in the "green" cones. The sensations produced by the spectrum will be predominantly blue from about 400 to 500 $m\mu$, and predominantly yellow from 500 to 700 $m\mu$, with a white or neutral point at about 500 $m\mu$. These neutral points will be the regions of greatest hue discrimination because of the great changes that the V curve on the one hand and both the G and R curves on the other show toward each other. An exactly analagous state of affairs would result if the substance in the blue-receiving cones were to become converted into either the "red" substance or the "green" substance. There would be a white spot, or neutral point at about 575 which would be the region of greatest hue

change. The information about the hue discrimination and neutral points of color blind individuals is entirely consistent with these deductions.

I am convinced, however, that color blindness cannot be completely described in such comparatively simple terms. What further ideas and limitations must be added it is not possible to say at this moment. But it is necessary to indicate where the weak spot is, so that further work and thought may be directed to it. It concerns the luminosity distribution of the spectrum for color blinds. We know from the excellent measurements of Brodhun[4] and others that the luminosity of the spectrum for the green blind, or deuteranope, does not differ significantly from the normal. Although adequate measurements are not available for the blue-blind or tritanope, it is probable that this type also has a practically normal distribution of brightness in the spectrum. Both these facts are readily understandable in terms of the V, G, and R curves. The curves are so similar that an exchange of one for another does not materially affect the luminosity distribution in the spectrum. Certainly it would be difficult to be sure of any such difference what with the large individual variations in the luminosity function and the relative uncertainty of heterochromic photometry.

The case of the red-blind or protanope is strikingly different. It has been known for years,[23] and the fact has been repeatedly verified,[9] that the luminosity distribution in the spectrum is very different for the protanope as compared with the normal or the deuteranope. The maximum is shifted toward the short-wave end of the spectrum. The luminosity is sharply depressed in the long-wave end, and it is correspondingly increased in the short-wave region. This is known relatively from the luminosity measurements themselves, and is corroborated by the enhanced visual acuity in the same region in proportion to its reduction in the long-wave end. The simple transformation of the R curve into the G curve, though it is in the right direction, is not adequate in magnitude. The resulting change in brightness is only about a fourth of what is necessary to describe the data obtained by Koenig,[23] by Exner[9] and by others.

I have been unable to suggest any additional changes which, when combined with the transformation of the "red" substance into the "green," would account for the singular character of the luminosity function of the red-blind, and at the same time remain consistent with all the other things that are known about color blindness. For example, the obvious suggestion that the V, G and R curves be separated more

than they are at present does not help matters. Their present separation is practically the maximum possible provided their areas are kept about equal. My judgment is that the V, G and R curves are already separated more than necessary. This is because I at first thought that their separation could be made to account for the red-blind luminosity. However, some other derivation will be required to adjust this difficulty, and I hope that one of the effects of the presentation of the ideas in this paper will be such a formulation by someone else.

XIV. BLUE AND VIOLET

It has been apparent that the spectral distribution curve of one of the three cone primaries has been labelled V. This is because violet seems to act as a unitary thing in color mixture, and in hue discrimination, and I wished to stick as close to the data as possible. Nevertheless, one cannot ignore the fact that blue, like yellow or red is a unique sensation, whereas violet is a color blend obviously compounded of blue and red. It is for this reason that I spoke in Section IX of a blue receiving cone rather than a violet receiving cone.

Koenig got out of this dilemma in the first construction of his Grundempfindungen[26] by calling the curve B and by assuming that the red sensation curve has a secondary maximum in the short wave part of the spectrum and crosses the green sensation curve in the same region. After his investigations with tritanopes or blue-blinds, Koenig[24] abandoned this idea because this type of color blind showed only one neutral point, at 575 $m\mu$, and failed to show a neutral point at the supposed crossing of the red and green curves in the violet end of the spectrum. He then chose his Grundempfindungen to be violet, green, and red, even though most people have ignored his later work and have quoted only his former.

The argument from the blue-blind neutral point is perfectly valid and it is for this reason that I have also chosen V as the designation of the spectral distribution of one of the primaries. Curiously enough, the most recent study of a blue-blind person, unfortunately reported only in a very sketchy form,[6] established a neutral point in the violet end of the spectrum. If this turns out to be the real condition, then Koenig's first choice of a blue Grundempfindung and a secondary crossing of the red and green curves will have to be reconsidered.

So far as the V, G, and R curves are concerned the situation is the same. At present the G and R curves cross but once. If the secondary neutral point of the tritanope becomes established, then it is a simple matter to make the G and R curves cross at about 420 $m\mu$ and to desig-

nate the third curve by B and not by V. No quantitative relations of any size are changed by such a procedure because of the relatively trivial alterations required in the R and G curves to accomplish this. In the meantime, however, one may keep the V, G, and R curves as they are. But that leaves us the dilemma of a V curve and a "blue" cone.

It is not a very difficult dilemma to resolve, and I give one possible solution here more for its interest than for my belief in it. Let us assume that all normal people are very slightly red blind. Suppose that the photosensitive substance in a small fraction, say 10 percent, of the red-receiving cones takes on the spectral characters of the substance in the blue-receiving cones. Then any light which normally affects the blue cones more than the red or green cones will also affect these red cones with the blue substance in them. This will be particularly the case at the short wave end of the spectrum. But the sensation which results from the stimulation of the originally-red-but-now-blue cones is still red. The result of this will be a constant association of a certain fraction of red sensation with the pure blue sensation, which gives the appearance of violet. Thus though violet is a blend it will behave like a unit in color mixture and in other quantitative aspects of color mixture.

This idea is perhaps capable of an experimental test. If the number of red, green, and blue cones are about equal in the fovea, and some of the red cones have taken on the properties of the blue, then the area under the resulting violet curve should be larger than that under either the green or the red; and the area under the red curve should be less than that under the green. Just how to ascertain this, I do not at present know. I, for one, hope that Dieter's discovery[6] of a secondary neutral point for blue-blinds is confirmed; that would facilitate and simplify the whole situation because then one could have a B curve, and a crossing of the R and G curves.

XV. FINAL CONSIDERATIONS

Reduced to its bleakest terms, what I have presented here is a series of relationships between well-known data. These data are often inadequate because they have been secured by various observers whose color systems may be sufficiently different to obscure their relations. Therefore I hope that no one will consider these ideas and computations as a final statement. Indeed, the situation is precisely the reverse. Experiments are exciting in proportion to the ideas which determine their existence. Many measurements and many experiments must be made before we can understand color vision. My present paper is thus largely a program. It represents the general ideas which are orienting us in our laboratory in an attack on the theoretical structure of color vision.

Thomas Young, one hundred and thirty years ago suggested the original idea for a theory of the mechanism of color vision. We have seen that this original idea is adequate as a foundation on which to build a theory. What I have tried to show is that when this idea is made precise and quantitative, and when there is added to it the suggestions of other workers, like Helmholtz, Koenig, and Fick, it is possible to account accurately and adequately for many previously obscure and unrelated but very important phenomena of color vision. Thomas Young's contribution to color vision is thus a beautiful leaf in the wreath of his scientific achievement. I hope that this leaf will continue to be green and vigorous for many years to come.

BIBLIOGRAPHY

[1] Abney, W. de W. Researches in color vision. London, 1913.

[2] Adrian, E. D. *Die Untersuchung der Sinnesorgane mit Hilfe elektrophysiologischer Methoden.* Ergebn. der Physiol., *26*, p. 501; 1928.

[3] Aubert, A. Physiologie der Netzhaut. Breslau, 1865.

[4] Brodhun, E. Beiträge zur Farbenlehre. Dissertation. Berlin, 1887.

[5] Brückner, A. *Zur Frage der Eichung von Farbensystemen.* Z. f. Sinnesphysiol., 1927, *58*, p. 322; 1927.

[6] Dieter, W. *Über die subjektiven Farbenempfindungen bei angeborenen Störungen des Farbensinnes.* Z. f. Sinnesphysiol., *58*, p. 73; 1927.

[7] Erlanger, J., and Gasser, H. S. *Analysis of the action potential in nerve.* The Harvey Lectures, Series 22, p. 90; 1926–1927.

[8] Exner, F. *Über die Grundempfindungen in Young-Helmholtz' schen Farbensystem.* Sitzungsber. Akad. Wiss., Wien, *111*, IIa, p. 857; 1902.

[9] ———. *Helligkeitsbestimmungen im protanopen Farbensystem.* Sitzungsber. Akad. Wiss., Wien, *130*, Abt. IIa, p. 355; 1921.

[10] Fick, A. Die Lehre von der Lichtempfindung. In L. Hermann, Handbuch der Physiologie, iii., Part 1, p. 139. Leipzig, 1879.

[11] Forsbes, A. The mechanism of reaction. In C. Murchison, The Foundations of Experimental Pychology, p. 128, Worcester, 1929.

[12] Hcht, S. *On the binocular fusion of colors and its relation to theories of color vision.* Proc. Nat. Aecad. Sci., *14*, p. 237; 1928.

[12a] Hecht, S. *The relation between visual acuity and illumination.* J. Gen Physiol., *11*, p. 255; 1928.

[13] Hiecke, R. *Neue Folgerungen aus den Farbenempfindungskurven von A. Koenig und C. Dit erici.* Z. f. Sinnesphysiol., *58*, p. 111; 1927.

[14] eHelmholtz, H. Handbuch der physiologischen optik. Hamburg and Leipzig, 1st edition, 1866.

[15] ———. Popular lectures on scientific subjects. Translated by E. Atkinson, London, 1881.

[16] ———. *Versuch das psychophysische Gesetz auf die Farbenunterschiede trichromatischer Augen anzuwenden.* Z. f. Sinnesphysiol., *3*, p. 517; 1891.

[17] Hering, E. Grundzüge der Lehre vom Lichtsinn. Berlin, 1920.

[18] Ives, H. E. *The transformation of color mixture equations from one system to another. II. Graphical aids.* J. Franklin Inst., p. 23; 1923.

[19] Jones, L. A. *The fundamental scale for pure hue and retinal sensibility to hue differences.* J.O.S.A., *1*, p. 63; 1917.

[20] Judd, D. B. *Chromatic visibility coefficients by the method of least squares.* J.O.S.A. & R.S.I., *10*, No. 6, p. 635; 1925.

[21] ———. *Sensibility to color change determined from the visual response functions; extension to complete and partial dichromasy.* J.O.S.A. & R.S.I., 1928, *16*, p. 115; 1928.

[22] Koenig, A. Ueber die neuere Entwickelung von Thomas Young's Farbentheorie. In A. Koenig, Gesammelte Abhandlungen zur Physiologischen Optik, Leipzig, p. 88, 1903.

[23] ———. Ueber den Helligkeitswerth der Spectralfarben bei verschiederer absoluter Intensität. In A. Koenig, Gesammelte Abhandlungen zur Physiologischen Optik, Leipzig, p. 144, 1903.

[24] ———. Ueber "Blaublindheit." Sitzungsber. Akad. Wissensch., Berlin, p. 718; 1897.

[25] Koenig, A., and Dieterici, C. *Ueber die Empfindlichkeit des normalen Auges für Wellenlängenunterschiede des Lichtes.* Wiedem. Ann. Physik u. Chem., *20*, p. 579; 1884.

[26] ———. *Die Grundempfindungen in normalen und anomalen Farbensystemen und ihre Intensitätsverteilung im Spektrum.* Z. Psychol. u. Physiol. Sinnesorgane, *4*, p. 241; 1892.

[27] v. Kries, J. Die Gesichtsempfindungen. In Nagel's Handbuch der Physiologie des Menschen, Braunschweig, *3*, p. 109; 1905.

[28] Ladd-Franklin, C. Colour and colour theories. New York, 1929.

[29] Laurens, H., and Hamilton, W. F. *The sensibility of the eye to differences in wave length.* Am. J. Physiol., *65*, p. 547; 1923.

[30] Maxwell, J. C. *On the theory of compound colors and the relations of colors of the spectrum.* Scientific Papers, Cambridge, *1*, p. 410; 1890.

[31] Mayer, A. M. *The history of Young's discovery of his theory of colors.* Am. J. Sci. and Arts, Series, 3, *9*, p. 251; 1875.

[32] Müller, J. Handbuch der Physiologie des Menschen. Coblenz, 1840.

[33] Newton, I. Opticks. Book 3, Queries, 13, 14, and 16. London, 1704.

[34] Peddie, W. Color Vision. London, 1922.

[34a] Priest, I. G. *A precision method for producing artificial daylight.* Phys. Rev., *9*, p. 502; 1918.

[35] Priest, I. G., and Brickwedde, F. G. *The minimum perceptible colorimetric purity as a function of dominant wave length with sunlight as neutral standard.* J.O.S.A. & R.S.I., *13*, p. 306; 1926.

[36] Rochat, G. F. *Étude de mélange binoculaire de rouge et de vert.* Arch. neérl. de Physiol., *10*, p. 448; 1925.

[37] Schrödinger, E. *Über das Verhältnis der Vierfarben- zur Dreifarbentheorie.* Sitzungsber. Akad. Wiss., Wien, *134*, Abt. IIa, p. 471; 1925.

[38] Sinden, R. H. *Studies based on spectral complementaries.* J.O.S.A. & R.S.I., *7*, p. 1123; 1923.

[39] Steindler, O. *Die Farbenempfindlichkeit des normalen und farbenblinden Auges.* Sitzungsber d. k. Akad. der Wiss., Wien, *115*, Abt IIa, p. 39; 1906.

[40] Trendelenburg, W. *Weitere Versuche über binokulare Mischung von Spektralfarben.* Arch. ges. Physiol., *201*, p. 235; 1923.

[41] Troland, L. T. *Report of the colorimetry committee of the Optical Society of America.* J.O.S.A., *6*, p. 527; 1922.

[42] Wright, W. D. *A re-determination of the trichromatic coefficients of the spectral colours.* Trans. Opt. Soc., *30*, No. 4, p. 141; 1928–29.

[43] Young, T. On the theory of light and colors. In Lectures in Natural Philosophy, London *2*, p. 613. 1807.

LABORATORY OF BIOPHYSICS,
COLUMBIA UNIVERSITY,
NEW YORK CITY.

ON THE REFRACTION OF A LIQUID BETWEEN TWO THIN PRISMS

By Morton Masius and W. E. Lawton

1. INTRODUCTION

The index of refraction of a liquid may be obtained by enclosing the liquid in a hollow prism, measuring the angle of the prism, A, and the minimum deviation of the transmitted light, D, and calculating the index of refraction of the liquid by the well known formula

$$n=\frac{\sin \tfrac{1}{2}(A+D)}{\sin \tfrac{1}{2}A}.$$

If great accuracy is desired various precautions are required, most of which are similar to those necessary in the case of a solid prism.* In addition to these requirements the simple theory also assumes that the glass covers of the prism are optical plates in the absolute sense of the word, *i.e.*, that they are bodies bounded by two plane and *parallel* surfaces. If the surfaces are not parallel the light passes through a triple prism, the two outside prisms of which are thin. We have been unable to find in the literature any treatment which answers the questions: (1) how closely parallel must the surfaces be, if the index of refraction calculated from the simple formula is to come within a certain accuracy, *e.g.* one unit of the fifth decimal; and (2), if the parallelism is not good enough, how may a correction for the prismatic shape of the covers be made. Since the hollow prism method is still widely used and, for absolute values, is probably as good or better than the other methods, the following brief investigation may be of interest to those pursuing such measurements.

2. GENERAL THEORY

The following symbols will be used: n the index of refraction of the liquid; g the index of refraction of the glass covers; A the total external angle of the triple prism; α and β the angles of the thin cover prisms; $A-(\alpha+\beta)$ the angle of the liquid prism; i, a, b, r, s, c, d, e the angles of the light rays with the respective normals taken in order. Thus i is the angle of incidence, e the angle of emergence, a and b are the angles in the thin prism α, r and s the angles in the liquid prism, and c and d the angles in the thin prism β. The total deviation of the light ray will be denoted by δ.

* See the article by J. Guild on Spectroscopes and Refractometers in R. Glazebrook's Dictionary of Applied Physics.

By the assumption that the index of refraction of air is unity, n and g become the indices of refraction of liquid and glass respectively with respect to air. The reduction to absolute indices may be made in the well known manner. Since, as will appear later, the value of g has very little influence on the result, it is superfluous to distinguish the values of g in the two prisms α and β.

In order to obtain a final answer which is correct in all cases, no matter what the relative direction of the angles α and β with respect to the angle of the liquid prism may be, it is well to adopt a general convention of signs for all angles. The sign rules used by J. Guild[1] were found to answer the purpose. They include all possible cases and are not in conflict with either the law of refraction or the necessary geometrical laws connecting the angles within any one prism, *e.g.*, a, b, and α.

We now have the following equations:

$$\left.\begin{aligned} \sin i &= g \sin a \\ \sin e &= g \sin d \\ g \sin b &= n \sin r \\ g \sin c &= n \sin s \end{aligned}\right\} \quad \text{I}$$

$$\left.\begin{aligned} \alpha &= a - b \\ \beta &= c - d \\ A - (\alpha + \beta) &= r - s \\ \delta &= A - i + e \end{aligned}\right\} \quad \text{II}$$

To these general equations we add the condition for minimum deviation:

$$\frac{\cos s \cos d}{\cos e \cos c} = \frac{\cos r \cos a}{\cos i \cos b}. \quad \text{III}$$

Hereafter it will be assumed that all variable angles are taken for the case of minimum deviation, and the symbols i, e, δ, etc. will be used accordingly without a distinguishing subscript. All equations are written in accord with our sign rule. Equation III is obtained by differentiation of the last equation of II with respect to any convenient variable, say r. If the derivative is equated to zero and this equation is combined with the derivatives of the other equations of I and II, equation III will result.

We may now eliminate the eight quantities i, a, b, r, s, c, d, e from

[1] l.c.

our nine equations and obtain a relation between n, g, A, α, β, and δ. This final equation in its general form is exceedingly complicated, and it is not possible to discuss and interpret it conveniently for our present purpose, which is to form an estimate of the influence of α and β on the value of n. Since we are chiefly interested in the case in which both α and β have small values, we shall proceed to obtain a much simpler expression which will answer our purpose by using certain approximations:

3. THE FIRST APPROXIMATION

For a quantity h which is small compared with unity and an angle ϵ which is small compared with ϕ we shall use the following approximations:

$$\frac{1}{1+h}=1-h \tag{1}$$

$$(1+h_1)(1+h_2)=1+h_1+h_2 \tag{2}$$

$$\sin\epsilon=\epsilon \text{ and } \cos\epsilon=1 \tag{3}$$

from which we find

$$\sin(\phi+\epsilon)=\sin\phi(1+\epsilon\cot\phi) \tag{4}$$

$$\cos(\phi+\epsilon)=\cos\phi(1-\epsilon\tan\phi). \tag{5}$$

If the cover prisms α and β are thin, the path of the light ray does not differ much from that in the case of cover plates, and the angles x and y defined by

$$x=r-\frac{A-\alpha-\beta}{2}=s+\frac{A-\alpha-\beta}{2} \tag{6}$$

and

$$y=i-\frac{A-\delta}{2}=e+\frac{A-\delta}{2} \tag{7}$$

are small angles. We shall also use an auxiliary angle z defined by the equation

$$g\sin z=\sin\frac{A-\delta}{2}. \tag{8}$$

The angle z may be thought of as a known angle. Using the approximations given, we may express the unknown angles a and d in terms of z as follows:

$$1+\cot^2 a=\frac{1}{\sin^2 a}=\frac{g^2}{\sin^2 i}=\frac{g^2}{\sin^2\left(\frac{A-\delta}{2}+y\right)}$$

$$=\frac{g^2}{\sin^2\frac{A-\delta}{2}}\left(1-2y\cot\frac{A-\delta}{2}\right)$$

$$=\frac{1}{\sin^2 z}-\frac{2y\cot\frac{A-\delta}{2}}{\sin^2 z}$$

$$=1+\cot^2 z\left(1-2y\frac{\cot\frac{A-\delta}{2}}{\cos^2 z}\right)$$

Hence

$$\cot a=\cot z\left(1-y\frac{\cot\frac{A-\delta}{2}}{\cos^2 z}\right) \tag{9}$$

and similarly

$$\cot d=-\cot z\left(1+y\frac{\cot\frac{A-\delta}{2}}{\cos^2 z}\right). \tag{10}$$

Now from equations I we find

$$2n=\frac{\sin i\sin b}{\sin r\sin a}+\frac{\sin e\sin c}{\sin s\sin d}$$

or

$$2n=\frac{\sin\left(\frac{A-\delta}{2}+y\right)\sin(a-\alpha)}{\sin\left(\frac{A}{2}-\frac{\alpha+\beta}{2}+x\right)\sin a}+\frac{\sin\left(\frac{A-\delta}{2}-y\right)\sin(d+\beta)}{\sin\left(\frac{A}{2}-\frac{\alpha+\beta}{2}-x\right)\sin d}.$$

By the use of the approximations (1), (2), and (4) this reduces to

$$2n=\frac{\sin\frac{A-\delta}{2}}{\sin\frac{A}{2}}\left[2-\alpha\cot a+\beta\cot d+(\alpha+\beta)\cot\frac{A}{2}\right]$$

and in view of (9) and (10) this yields

$$n=\frac{\sin\frac{A-\delta}{2}}{\sin\frac{A}{2}}\left[1-\left(\frac{\alpha+\beta}{2}\right)\left(\cot z-\cot\frac{A}{2}\right)\right] \tag{11}$$

which is our final result.

4. CLOSER APPROXIMATIONS

Before we draw further conclusions from equation (11) we may consider the following three points: (1) the effect of the omitted terms of higher order, (2) the order of magnitude of x and y, (3) the case in which α and β are nearly equal and of opposite sign.

1. In view of the extreme smallness of the small angles in the case of prisms the covers of which are nearly plates it is not necessary to justify the omission of the terms containing the second and higher powers of our small angles in any *one* step of the calculation. Since, however, our calculation contains many steps involving such omissions, there might exist some doubt regarding a perceptible cumulative effect of such omissions in *many* successive steps. To refute any such objection, we have recalculated n, using processes similar to those that lead to equation (11) but retaining both the first and the second powers of all small quantities. Since these calculations are very extensive and of mathematical rather than physical interest, we give merely the final answer. It is

$$n=\frac{\sin\frac{A-\delta}{2}}{\sin\frac{A}{2}}\left[1-\left(\frac{\alpha+\beta}{2}\right)\left(\cot z-\cot\frac{A}{2}\right)-\left(\frac{\alpha+\beta}{2}\right)^2\cot\frac{A}{2}\left(\cot z-\cot\frac{A}{2}\right)\right.$$
$$\left.+\frac{1}{2}\left(\frac{\beta-\alpha}{2}\right)^2\frac{\left(\sin^2\frac{A-\delta}{2}-\sin^2 z\right)\left(\sin^2\frac{A}{2}-\sin^2 z\right)}{\left(\sin^2\frac{A-\delta}{2}-\sin^2\frac{A}{2}\right)\sin^2 z\cos^2 z}\right] \tag{12}$$

or

$$n=\frac{\sin\frac{A-\delta}{2}}{\sin\frac{A}{2}}\left[1-\left(\frac{\alpha+\beta}{2}\right)\left(\cot z-\cot\frac{A}{2}\right)-\left(\frac{\alpha+\beta}{2}\right)^2\cot\frac{A}{2}\left(\cot z-\cot\frac{A}{2}\right)\right.$$
$$\left.+\frac{1}{2}\left(\frac{\beta-\alpha}{2}\right)^2\frac{(g^2-1)(g^2-n^2)}{(n^2-1)g^2\cos^2 z}\right] \tag{13}$$

These last equations show the following properties:

If $$\alpha=\beta=0, \quad n=\frac{\sin\frac{A-\delta}{2}}{\sin\frac{A}{2}}.$$

If $\alpha=\beta\neq 0$, the last term drops out.

If $$z=\frac{A}{2}, \quad n=\frac{\sin\frac{A-\delta}{2}}{\sin\frac{A}{2}} \text{ and } g=n.$$

If $z=(A-\delta)/2$, the last term drops out, $g=1$, and

$$n=\frac{\sin\left(\frac{A-\delta}{2}-\frac{\alpha+\beta}{2}\right)}{\sin\left(\frac{A}{2}-\frac{\alpha+\beta}{2}\right)}.$$

All of these relations are in accord with the physical properties of these simple cases.

It is obvious that for small values of α and β, unless α and β happen to be practically equal and opposite which case is considered below, the terms depending on the squares of α and β contribute no appreciable amount to the value of n. The simpler equation (11) is, therefore, sufficient for practical purposes.

2. In establishing equation (12) it was necessary to calculate x and y, because they do not disappear from the second order approximation as they did from the first. To find x and y we expand

$$\frac{\sin i \sin b}{-\sin e \sin a}=\frac{\sin r \sin c}{-\sin s \sin d}$$

and

$$\frac{[\cos i \cos b}{\cos e \cos a}=\frac{[\cos r \cos c]}{\cos s \cos d}$$

and obtain

$$2y \cot\frac{A-\delta}{2}-\alpha \cot a=2x \cot\frac{A}{2}+\beta \cot d \tag{14}$$

$$2y \tan\frac{A-\delta}{2}-\alpha \tan a=2x \tan\frac{A}{2}+\beta \tan d \tag{15}$$

from which x and y are obtained with sufficient accuracy.[2] They are of the order of magnitude of $(\beta-\alpha)$.

3. In the case in which α and β are nearly equal and of opposite sign, the linear correction term depending on $(\alpha+\beta)$ becomes small, the term depending on $(\alpha+\beta)^2$ vanishes, and the term depending on $(\beta-\alpha)^2$ may be the decisive correction term. It is possible that even the second order approximation is not exact enough for this case. It is, however, not worth while to consider this case further, because the value of n certainly comes very near to the ideal value (with plate covers) and with small values of α and β the correction may be omitted entirely.

5. CONCLUSION

Returning now to equation (11) we proceed to calculate the effect of α and β on n.

Assuming A as approximately 60°, and taking n as about 1.25 and g as about 1.70, we find $(\cot z-\cot \tfrac{1}{2}A)$ to be about 0.8, while for $n=1.40$ and $g=1.58$ the factor $(\cot z-\cot \tfrac{1}{2}A)$ is about 0.3. This means then that if we calculate n from the simple expression

$$n=\frac{\sin\frac{A-\delta}{2}}{\sin\frac{A}{2}}$$

and make no allowance for α and β the value of n will be wrong by an amount (depending on g and n) equal to from one-half to one-fifth of the sum of α and β. If, therefore, this sum does not exceed one second, n cannot be wrong by more than 3 units of the sixth decimal, which, with the accuracy obtainable with our present day experimental equipment, is a *negligible* source of error.

If, on the other hand, the sum of α and β reaches the order of magnitude of one minute, the neglected correction completely vitiates the fifth decimal and probably makes the fourth decimal wrong by one or two units. We believe that this is actually the case in some of the older work on refractive indices of liquids.

[2] It is highly probable though perhaps not evident that x and y are of the same order of magnitude as α and β. We have, therefore, replaced equations (14) and (15) by the corresponding equations which contain not only the first but also the second power terms in x, y, α, and β. A careful consideration of these equations has left us entirely satisfied that the linear equations (14) and (15) give adequate values of x and y, and that nothing is to be gained from the use of the corresponding complicated quadratic equations.

If $(\alpha+\beta)$ is larger than one second, and if α and β can be measured or estimated, the value of n may still be obtained with the desired accuracy by the use of equation (11).

WORCESTER POLYTECHNIC INSTITUTE,
WORCESTER, MASS.,
DECEMBER, 1929.

THE AMOUNT OF ULTRAVIOLET RADIATION NECESSARY TO CURE RICKETS

By George H. Maughan and J. A. Dye

There can be no doubt that ultraviolet radiations, both from sunshine and artificial sources such as the quartz mercury arc and certain special carbon arcs, have great value in the prevention and cure of disease. It seems to the writer that prevention will ultimately prove to be the more important. Up to the present, however, we have much less data with regard to its power to prevent than to cure disease. This is true, first, because it is more difficult to estimate the amount of prevention in a group of well individuals than the degree of cure among those that have been ill. Second, we are as yet, unfortunately, more interested in curing disease than preventing it.

In the experimental field, data obtained from the irradiation of normal animals by ultraviolet are often misinterpreted and many times lead to confusion because beneficial treatments can at best only keep the animal normal. If doses are massive enough to upset equilibrium, the body has a remarkably strong tendency, and great potential ability to quickly regain its normal balance.

From a biological point of view, much remains to be learned about the action of ultraviolet rays on specific organs of the body. We do not know very definitely how deeply the rays penetrate into the tissues, nor even the minimum depth to which they must penetrate before they become useful. The particular wave lengths or combination of wave lengths and the time needed for irradiation are still problems in the study of which we have just made a beginning. It is with the length of exposure necessary to cure rickets in chickens that this paper is principally concerned.

The greatest handicap in the way of solving many of our problems is the lack of suitable instruments for the production, segregation and measurement of specific wave lengths. It must be evident that the field requires the co-operation of the Physicist and the Biologist.

WAVE LENGTHS HAVING THERAPEUTIC VALUE

The specific rays which affect the animal tissues and stimulate normal reactions, or the rays which cause more rapid repair after injury by disease, are not known. In some conditions response may not be due to any small band of radiations, but rather to a general effect from

the broad band of radiant energy to which life has become accustomed, and upon which it has, perhaps, become even dependent during the ages of its evolution. Such a disease as tuberculosis seems to respond better to sunlight than the radiations from artificial sources. Rickets, on the other hand, responds to treatments from a much more limited field of radiation; and artificial sources of the rays are very satisfactory.

In 1923 Hess and Weinstock (1) found that wave lengths longer than 3240A have no curative properties for rickets. In later experiments these authors, as well as a number of other investigators, have found that the upper limit of antirachitic rays is at about 3130A. It is more difficult to find the lower limit of valuable rays. Hess and Anderson (2) believe that with a mercury vapor lamp, wave lengths shorter than 2900A have greater power to cure rickets in rats, for the same number of ergs of energy which fall upon a given surface, than those between 2900A and 3660A. However, when ergosterol was irradiated by wave length 2800A healing was not so good as when the same energy from line 3025A was used. Our own experiments (3) showed that line 3025A in the mercury vapor spectrum, had some power in curing rickets when applied directly to chickens and that perhaps wave length 2894A had benefit, but there was no doubt that line 2968 was near the center of a relatively narrow band of antirachitic rays. It seemed that this line in the mercury vapor spectrum was more potent than all the others put together; and that the antirachitic power of other radiations was about equally divided between those just longer and those just shorter than line 2968A.

The irradiation of ergosterol involves quite a different condition, and rays most effective in making it antirachitic may be somewhat different from those which cure rickets in the irradiated animal.

With ergosterol we do not have a variety of tissues and pigments with which to deal. It appears from the work of Hilborn (4) and his co-works and also from that of Rosenheim and Webster (5) that wave lengths shorter than 270mμ destroy vitamin D, and that rays at about 280mμ are most readily absorbed and presumably those which cause the greatest effect in changing ergosterol into vitamin D. The absorption bands which appear at 270mμ and 293.5mμ are well defined but not so strong as the one at 280mμ.

DEPTH OF PENETRATION

Until recently it was generally accepted that ultraviolet shorter than 3000A did not penetrate the skin deeper than 0.1 mm. In 1928 Macht,

Anderson and Bell (6) reported experiments in which they concluded that "penetration of ultraviolet rays through living skin and other tissues is much greater than has hitherto been supposed," and that there is a marked difference between living and dead tissue (in favor of living tissue). They also thought that the shorter waves (2800A) penetrated more deeply than those above 2900A. This paper brought a prompt reply from Leonard Hill (7) who pointed out that tuberculosis bacilli are not killed in one hour when rotated in serum under a mercury vapor arc; while they are easily killed in Ringer's solution. To destroy paramecia it requires about six times the killing dose when a screen of omentum is interposed.

Bachem (8) also showed that errors are evident in the work of Macht, Anderson and Bell, "and that while absorption of light by proteins decreases slightly near 2500A and 2600A before the end absorption sets in, the decrease from 3700A to 2800A is absolutely out of the question." Anderson and Macht (9) repeated the earlier experiments of Macht, Anderson and Bell and found that between six and ten percent of the short ultraviolet penetrate through a thickness of 1.2 mm of rabbit's skin. It seems that depth of penetration is greater than was at first accepted but that much of this radiation is stopped by superficial cell layers or even the lymph which bathes the cells or the cytoplasm of the cells themselves.

THE SKIN

Along with the problem of penetration of health promoting rays into the skin, comes the question of the thickness of this very important organ. Careful histological studies and measurements show that the total thickness varies between .03 and 4.0 mm, being thickest on the palms of the hands and soles of the feet; and thinnest in the regions of the body where there is least friction, such as the inside surface of the arms.

The average thickness is about 1.0 mm. The epidermis, or outside part which is made up of several distinct layers of cells, has a normal thickness of 0.03 to 0.5 mm. There are no blood vessels in the epidermis but in its deeper layers we find the pigment of the skin. It seems that the light rays must penetrate the epidermis, and reach the blood and nerves just beneath, in order that benefit may be derived from them. The germ layer of the epidermis which is the dividing line between it and the dermis is very irregular in contour. Small finger like projections from the dermis push this stratum germinativum upward into millions

of little finger like papillae from .03 to 0.2 mm in height. Into some of these papillae project the endings of nerves and into others loops of blood capillaries. Radiations which are capable of penetrating the epidermis to the tips of these papillae are no doubt absorbed, in a large part, by the blood, and the products which result are quickly carried away to other parts of the body.

Rollier, (10) the greatest of all advocates for the use of sunshine as an aid in the cure of tuberculosis, says that "the skin which has important functions in aiding general metabolism, elimination, immunization and internal secretion, is kept in a high state of health by the radiations from the sun." It seems reasonable that such a skin in a high state of health would be less permeable to microörganisms; it certainly has a better blood supply and responds more quickly and more perfectly to outside influences such as heat and cold.

THE AMOUNTS OF EXPOSURE NECESSARY IN THE CONTROL OF RICKETS

There is not much definite evidence as to the length of exposure required to bring about the best results in the use of ultraviolet. It is likely that different diseases require different exposures. We believe that in much of the work done in the past, longer irradiations have been given than are necessary. There is no doubt that frequently over exposure rather than being helpful is actually harmful to the body. This applies particularly in the treatment of human beings and when artificial sources such as the mercury vapor arc are used.

Rickets is a disease which responds so promptly and regularly to ultraviolet radiations that results, following exposure to a source of known strength, can be predicted. Recovery is so constant that healing of the disease has the regularity of a chemical reaction. Hess and Weinstock (2) found that two minutes exposure to a quartz mercury vapor arc at a distance of three feet and with 75 volts would prevent rickets in rats.

Goodale (11) showed that ten minutes treatment per day prevented rickets in chickens at distances from three to seventy inches and that one minute per day was as adequate as the longer exposures. Heuser and Norris (12) of the Poultry Department of Cornell University found that 10.9 minutes exposure per day to Ithaca sunshine during the months of July and August prevented rickets in chickens. It required somewhat longer treatments, 13.7 minutes, from the Uviarc Poultry treater (an ac 110 volts quartz mercury arc) to bring about the same results. More recent work (not published) under the direction of

Dr. L. C. Norris shows that much less time than ten minutes will prevent rickets in chickens when exposed to noon-day sunshine. This included the radiations on cloudy days. The skyshine on fair days, but when the sun is hidden by clouds, is considerable. Shrader, Coblentz and Korff (13) have found that seventy percent of the ultraviolet from solar radiation at Baltimore is in skyshine.

Dr. Edwin T. Wyman (personal communication) of the Children's Hospital in Boston has found that children suffering from rickets heal more rapidly and recover more quickly than rats having the disease and treated by the same source of light with the same length of exposure. Gerstenberger and Hartman (14) attempted to find "the minimal dose of ultraviolet rays as produced by the quartz lamp which with exposures given once weekly would cure rickets within the customary time period." Their results showed that twenty minutes exposure for negro children, and a total of ten minutes, for whites gave the required results. The distance was 80 cm from the lamp to the patient.

Experiments on the prevention of colds by regular weekly irradiations of Freshmen and Sophomore students at Cornell University (15), (16) resulted in a reduction in frequency and also severity of the colds. Ten minutes exposure per week at a lamp-table distance of 30 inches gave an average decrease of 27.9% to 55.0%.

EXPERIMENTAL

In the present work chickens were used as experimental animals. Chickens have several advantages, such as uniformity of breeding and uniformity of age; the food given from the very day of birth is the same and there is a uniformity of response in the development of rickets when ultraviolet radiations or some special source of Vitamin D is withheld. These conditions make them valuable animals in studying irradiation problems.

In the stock which we have used, rickets begins to appear at about 24 days after hatching; and at six weeks 85% of the animals have a severe form of the disease. Most of the remaining 15% may not develop severe symptoms for several weeks or not at all. On a balanced diet consisting of a mash mixture containing 20% wheat bran, 10% wheat flour middlings, 29% yellow corn, 10% dried buttermilk, 10% ground chopped oats, 10% meat scrap, 10% steamed bone meal, and 1% salt, results are about as uniform as when a more limited diet of 57% yellow corn, 20% dried skim milk, 20% wheat middlings, 2½% ground bone and ½% salt, is used. It is evident that ordinary foods do not have any

marked effect in the prevention of rickets. Likewise, rickets is about as severe and the time when the disease appears about as early when the more complete ration is fed.

The method of estimating severity of rickets is important in a quantitative study of the problem, especially when at the beginning of irradiations a uniform severity of the disease is required, and when at the end less than a complete cure is desired. Our estimates are based upon a study of the general appearance of individual animals, growth curves and x-ray photographs taken at the beginning of irradiation, at the end of two weeks, and at the close of the experiment. See Plate I. In addition post mortem examinations and also blood and bone analyses were made and studied for each group. Wherever percent recovery is given in the following tables, the figures have been based upon such a composite analysis. While any one of these six studies gives a fair idea of the condition of the animal and the severity of rickets, and while these checks run relatively parallel during the progress of the disease, still they are all essential in making an estimate of the true condition.

In 1926 chickens from two series of experiments were cured of rickets in twenty-eight days with seven minutes per day exposure to a quartz mercury arc. A Uviarc Poultry treater with a lamp-floor distance of 30 inches on the ordinary ac 110 volt electrical supply furnished the radiation. No attempt was made to measure the voltage or current across the arc. In the first series, the animals were four weeks old, and in the second five weeks old when irradiations began. In another experiment of the same year chickens were given ten minutes exposure under identical conditions except that they were considerably older (ten weeks) and had suffered from rickets longer. In a fourth series five minutes treatment was given to seven weeks old chickens. The group receiving ten minutes irradiation did not respond quite so rapidly as the younger animals which received seven minutes treatment, and the animals receiving five minutes irradiation appeared not to have sufficient for optimum recovery.

Typical results from this year's experiments (1926) are given in Table 1.

Attention is called here to Groups 1 and 3. Group 1 averaged 64 grams increase in weight in 28 days while group 3 averaged 244 grams gain in the same time. The actual total exposure was only 147 minutes.

Experiments conducted in 1927 showed that severe rickets could be 95% cured in 28 days with one minute per day exposure. Here the lamp-floor distance was 22½ inches and the electrical supply did not

fluctuate. This current was generated in the laboratory and was held at 110 volts on the line (80 volts across the arc). The lamp was a dc Copper Hewitt tube in a special housing which prevented reflection. One-half minute irradiation per day under the above conditions brought about very noticeable improvement (67% cure). Ten minutes exposure per day cured rickets more completely than one minute, but the increased benefit came in the more rapid acceleration of recovery during the first week of treatment.

TABLE 1. *Growth records.*
(1926 Experiments)

Group number	Number of chicks	Average weight in grams				Average condition	
		1st week	2nd week	3rd week	4th week	Beginning	End
(1) No Irradiation	11	261	317	320	325	Rachitic	Very rachitic
(2) Heads & Feet Irrad. (7 minutes)	11	245	364	400	491	Very rachitic	Normal
(3) Total Irradiation (7 min.)	11	228	350	388	472	Very rachitic	Normal

In 1929 two groups of chickens were given one minute twenty seconds exposure to the same kind of tube in the same reflector as was used in 1927. In this work the distance was 30 inches from the floor of the cage to the lamp. The current was the same as in 1927 and held constant at 115 volts on the line and 75 volts, 4½ amperes across the burner. Cure was 83% with the rays coming from above and 94% complete recovery with the rays from the side of the animals. The advantage of the irradiations from the side was no doubt due to less shading of the legs by the chicken's own body. In this series, five minutes daily exposure was more than an optimum dose. Two minutes would have caused 95% cure in twenty-eight days. As a comparison with the two groups having an irradiation time of one minute and twenty seconds for twenty-eight days, another series of this year contained a group of chickens where all of the radiation from ultraviolet was received in four exposures. The treatments were given on the first day of each of the four weeks of the test. The total time of exposure

TABLE 2. *Growth records.*
(1927 Experiments)
Average weights (gms) at six weeks and at the end of each week for 28 days treatment.

Pen No.	Treatment	Begin.	1st week	2nd week	3rd week	4th week	Condition	
							Begin.	End
1.	no irrad.	263	281	294	292	291	severe rickets	very severe
2.	21¾ min. window glass	275	307	326	336	335	severe rickets	very severe
6.	one half min.	273	282	328	383	449	severe rickets	3 normal 3 rickets
7.	1 min.	256	280	353	439	526	severe rickets	7 normal 2 mild rickets
10.	10 min.	261	315	386	509	596	severe rickets	10 normal

TABLE 3. *Results of blood and bone analyses.*
(1927)

Pen No.	Treatment	Ca	P	Ca×P*	Bone ash	Recovery
1	No irrad.	7.65	2.76	21.11	9.15	None
2	21¾ min. window glass	7.52	2.78	21.00	9.61	None
6	30 sec. open arc.	8.51	4.30	37.55	14.46	67%
7	60 sec. open arc.	10.30	4.82	49.24	17.53	95%
10	600 sec. open arc.	10.56	4.60	48.30	16.29	100%
11	sunlight	11.35	5.71	64.41	15.41	100%

* The product obtained by multiplying the units of calcium by the units of phosphorus.
Note: Sunshine group had outdoor range and access to green food and insects.

was 37½ minutes. This was the same as the one minute twenty second groups had received. Recovery was nearly complete in this group. The male birds were practically cured and the females averaged 90% recovery. This experiment proved that the effects of irradiation carries over and keeps the stimulus for growth and recovery going for at least one week. This was also noted in the 1926 experiments where irradiations were discontinued for seven days, beginning at the end of one week's irradiation, Table 1. It is also clear that relatively massive doses can be given to take the place of short but more frequent exposures.

TABLE 4. *Growth records.*
(Series IX. 1929)
Showing average growth of chickens and recovery from rickets while receiving four weekly irradiations, (total 37½ min.) in the four weeks experimental period.
Pen 17—non-irradiated controls.

Pen	Average weight in grams					Condition beginning	Recovery end
	Begin. week	1st week	2nd week	3rd week	4th week		
16 Males	393	447	529	659	738	Severe rickets	95%
16 Females	316	322	339	398	510	Severe rickets	90%
17 Males	402	431	457	467	457	Severe rickets	Very severe rickets
17 Females	327	346	363	378	372	Severe rickets	Very severe rickets

We have often noted that the male chickens (White Leghorn) which have reached the age of six weeks recover more quickly, or with less exposure to ultraviolet rays, than do the females. It is evident on analysis of the data at hand that this is not due to a difference in the power of the sexes to react to irradiations, but rather to the fact that the males are developing combs by this age while those of the females are still rudimentary. The comb is a highly vascular appendage and situated on top of the head it has a particular advantage for irradiation because of its position. As an actual minimum exposure for a 95% cure of rickets for male chickens under the conditions of the 1929 experiments, one and one third minutes daily is required.

To accomplish the same results with females two and two thirds minutes are necessary. The amount of irradiation for females is approximately double that needed for males. Table 5 gives the growth record which bears this out, note especially Pen 3. In a number of such records as these this fact is made clear.

TABLE 5. *Growth records.*
(1929 Experiments)
(Rickets prevented in Pen 1 by ten minutes irradiation alternate days. Cure of rickets Pens 3, 8 and 14, daily irradiations. Average weights in grams.)

Pen No.	Sex	Irradiation	Begin.	1st week	2nd week	3rd week	4th week	Percent recovery
1	Male	10 minutes	345	447	554	678	788	100%
1	Female	10 minutes	339	427	530	612	734	100%
2	Male	None	240	248	255	261	289	0%
2	Female	None	221	226	228	239	240	0%
3	Male	1 min. 20 sec.	207	241	305	387	500	95%
3	Female	1 min. 20 sec.	216	237	269	315	387	70%
8	Male	5 minutes	213	271	339	484	579	100%
8	Female	5 minutes	186	217	270	393	497	100%
14	Male*	1 min. 20 sec.	268	321	405	501	607	96%
14	Female*	1 min. 20 sec.	195	227	293	375	480	90%

* Radiation from the side.

It should be noted that growth is not proportional to exposure, and the curve cannot, therefore, be represented as a straight line, but is rather one with a lag at the beginning and in the section representing the amount of irradiation received in two or three minutes from the quartz mercury arc it is almost horizontal. In our earlier (3) experiments one minute daily exposure gave 95% cure while one half minute gave 67% cure.

Critical Area. Another set of interesting experiments deals with the critical area of the body, that is, the part of the body through which the rays penetrate to the blood and other active tissues. In 1926 two series of experiments contained groups of chickens intended to throw light on this subject. In some cases the eyes were exposed and in others the heads and legs, and in still others the bodies were exposed to the radiations. The covering was a moderately heavy cotton or felt cloth. A lamp-floor distance of 30 inches and a line voltage of 110, a new Uviarc

Poultry Treater produced no benefit when the eyes alone were exposed. This was true where the cure of rickets was attempted in series I, with seven minutes exposure. In series II treatment was given for the prevention of rickets. When the bodies and legs were exposed cure of rickets or prevention of the disease seemed as complete as when no covering was applied. When the heads and legs were exposed for seven minutes, see table 1, group 1 and 2, recovery was as rapid as when the whole body was irradiated, group 3. This suggested that the legs must play a very important rôle in absorbing the rays.

In 1927 the critical area was studied in one pen of chickens. Those animals which had the bodies exposed with the heads and feet covered showed no improvement in growth above the non-irradiated controls. Those which had the feet, from the hocks down, exposed were cured in 28 days. The irradiation time was 3½ minutes daily, the lamp-floor

TABLE 6. *Area of absorption.*
(1927 Exp.)

	Average Weights (grams)					Condition	
	Begin.	1st week	2nd week	3rd week	4th week	Begin.	End
Body Exposed	300	295	313	299	299	Rickets	Severe rickets
Feet Exposed	241	283	364	463	561	Severe rickets	Normal

distance 22½ inches with 80 volts across the arc of the burner. The source of the rays was a new dc Uviarc quartz mercury arc on a 110 volt line. It is evident that for short exposures the feathers of White Leghorn chickens six to ten weeks old stop most, if not all of the antirachitic rays from a high power quartz arc at close range. This means that the feathers are like clothing in preventing the rays from reaching the body. The recovery of the animals having one side of the feet irradiated was about as rapid as that of the irradiated controls having five minutes exposure. Growth records of these animals are shown in Table 6.

We have just completed another series of tests where the above data have again been checked. In the 1929 work the lamp was a new dc burner. The line voltage and the voltage and the amperage across the arc were 115, 77 and 4½ respectively. The housing of the lamp had no reflectance. The burner was 30 inches above the floor of the irradiation

cage, and the exposure time was five minutes per day. Male animals having the heads only exposed (these chickens had combs at the beginning of irradiation, $\frac{1}{2} \times 1\frac{1}{2}$ inches) were 90% cured in 28 days. Female chickens, with only rudimentary combs, having the heads exposed showed only slight improvement over non-irradiated birds. There was no noticable benefit except a small increase in blood Ca and a slightly higher per cent of bone ash.

TABLE 7. *Growth records.*
(Average weights in grams)
(1929 experiments.)
Area of absorption

Pen No.	Sex	Treatment	Begin.	1st week	2nd week	3rd week	4th week	Percent recovery
3	Male	1½ min. total irrad.	207	241	305	387	500	95%
3	Female	1½ min. total irrad.	216	237	269	315	389	70%
8	Male	5 min. irrad.	213	271	339	484	579	100%
8	Female	5 min. irrad.	186	217	270	393	497	100%
13 A	Male	5 min. heads exposed	194	218	288	383	462	95%
13 A	Female	5 min. heads exposed	153	171	190	221	224	10%
13 B	Female	5 min. bodies exposed	167	181	197	236	248	15%
13 C	Female	5 min. feet exposed	149	185	273	374	455	100%

Chickens having the bodies exposed with feet and heads covered were only slightly benefited. The improvement might easily be accounted for by the fact that bare spots under the wings received accidental exposure while the chickens were under the lamp.

The third group with the unfeathered parts of the legs irradiated were cured of the severest type of rickets in 28 days. Their recovery was fully as rapid as that of the chickens having five minutes total irradiation daily. Growth records are given in table 7 and analysis of blood and bone in table 8, pen 13, A, B and C.

Two important facts stand out prominently as a result of the experimental work here reported. First, the total amount of ultraviolet necessary to cure or prevent rickets is extremely small when measured in calories, and second, if judged by their vital biological effects, certain rays have immence power, but they must fall upon the uncovered skin in order to reach the organism and be of value. The natural covering of

the body (feathers in these experiments) is a very definite block to ultraviolet rays. However, a fraction of the total area is capable of absorbing sufficient of the rays for the maintanence of normal health.

TABLE 8. *Results of blood and bone analysis.*
Area of absorption
(1929 experiments)

Pen	Treatment	Blood Ca Mgs per 100 cc. serum	Bone Ash % green wt.	Percent Recovery*
2	No irradiation	7.64	10.09	None
2	No irradiation	7.64	12.70	None
3	1 min. 20 sec. exp.	10.96	18.28	95
3	1 min. 20 sec. exp.	10.30	17.16	70
8	5 min. exposure	11.58	18.25	100
8	5 min. exposure	10.77	18.00	100
13	Heads exposed 5 min.	11.58	18.27	95
A	Heads exposed 5 min.	9.26	13.77	15
B	Bodies exposed 5 min.	9.26	13.36	10
C	Feet exposed	11.46	17.82	100

* The results of a study of general appearance, growth records, x-ray photographs, post mortem, blood Ca, and bone ash.

SUMMARY

1. Ultraviolet radiations are valuable in the cure and prevention of disease. The latter will perhaps prove to be the more important.

2. The rays having therapeutic value may not be exactly the same for different diseases, but for rickets they lie in the ultraviolet near the line 2968A.

3. A certain percent of these ultraviolet rays penetrate the epidermis and reach the blood capillaries in the papillae of the dermis .2 to .5 mm beneath the surface.

4. Less exposure is required to prevent rickets than to cure the disease.

5. In these experiments it was shown that one minute daily irradiation at a distance of 22½ inches and under the conditions specified, would cause a 95% cure of rickets. One half minute daily irradiation will induce 67% cure in 28 days. When ten minutes irradiation was given only slightly better results were obtained than for one minute.

One and one third minutes exposure at a distance of 30 inches caused 83% recovery from severe rickets. When the rays came from the side there was a 90% recovery. The better gains were due to less shading of the feet by the animal's own body. Four exposures of less than ten minutes each, given weekly caused as satisfactory recovery as the same total exposure (thirty seven and one third minutes) given in 28 doses of one and one third minutes each. Six weeks old male chickens, because of their combs, require less exposure to ultraviolet rays than female.

6. Feathers so completely obstruct the beneficial rays that the minimum adequate exposure is almost ineffective when the legs and heads are covered and the bodies are exposed.

DEPARTMENT OF PHYSIOLOGY,
CORNELL UNIVERSITY, ITHACA, N. Y.

BIBLIOGRAPHY

(1) Hess, A. F. and M. Weinstock. *Study of Light Waves in their Relation to Rickets.* Journ. Amer. Med. Assoc., *80*, p. 687; 1923.

(2) Hess, A. F., Wm. T. Anderson Jr. *The Antirachitic Activity of Monochromatic and Regional Ultra-Violet Radiation.* Journ. Amer. Med. Assoc., *89*, p. 1222; Oct. 1927.

(3) Maughan, George H. *Ultra-Violet Wave Lengths Valuable in the Cure of Rickets in Chickens.* Amer. Journ. of Physiol., *87*, 2, p. 381; Dec. 1928.

(4) Heilborn, I. M., E. D. Kamm and R. A. Morton. *The Absorption Spectrum of Cholestrol and Its Biological Significance with Reference to Vitamin D.* Biochem. Journ., *21*, p. 78; 1927.

(5) Rosenheim, O., and T. A. Webster. *Specificity of Ergosterol as Parent Substance of Vitamin D.* Biochem. Journ., *22*, p. 1426; 1928.

(6) Macht, David I., Wm. T. Anderson Jr., and F. K. Bell. *The Penetration of Ultra-Violet Rays Into Live Animal Tissue,* Journ. Amer. Med. Assoc., *90*, 3, p. 161; Jan. 1928.

(7) Hill, Leonard. *Penetration of Ultra-Violet Rays of the Mercury Vapor Lamp.* Journ. Amer. Med. Assoc., *90*, 16, p. 10; April 1928.

(8) Bachem, Albert. *The Penetration of Ultra-Violet Rays into Animal Tissue.* Journ. Amer. Med. Assoc., *90*, 7, p. 563; Feb. 1928.

(9) Anderson, Wm. T. Jr., and David I. Macht. *Penetration of Ultra-Violet Rays into Living Animal Tissue.* Amer. Journ. Physiol., *86*, p. 320; 1928.

(10) Rollier, A. *The Therapeutic Preventive and Social Value of Heliotherapy in Surgical Tuberculosis.* Journ. of State Medicine, *38*, 8, p. 435; Aug. 1928.

(11) Goodale, H. D. *Early Growth of Chickens with Special Reference to Ultra-Violet Light.* Amer. Journ. Physiol., *79*, p. 44; 1926.

(12) Heuser, G. F. and L. C. Norris. *Rickets in Chickens.* Poultry Science, *8*, 2, p. 89; 1929.

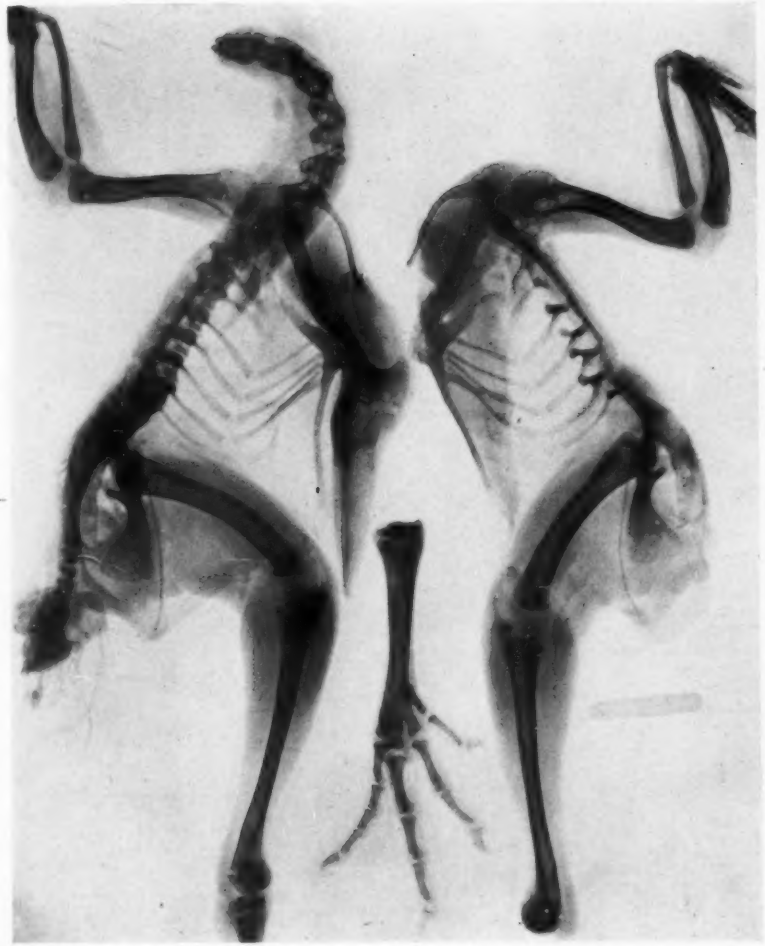
(13) Shrader, J. H., M. H. Coblentz and F. A. Korff. *Effect of Atmospheric Pollution Upon Incidence of Solar Ultra-Violet.* Amer. Journ. of Public Health. *19*, 7, p. 717; 1929.

(14) Gerstenberger, Henry J., and J. I. Hartman. *Quartz Lamp Therapy in Human Rickets.* Journ. Amer. Med. Assoc., *92*, 5, p. 367; Feb. 1929.

(15) Maughan, G. H., and Dean F. Smiley. *Effect of General Irradiation with Ultra-Violet Light Upon the Frequency of Colds.* Journ. Prev. Med., *2*, 1, 69; Jan. 1928.

(16) Maughan, G. H., and Dean F. Smiley. *Irradiations From a Quartz Mercury Vapor Lamp as a Factor in the Control of Common Colds.* Amer. Journ. of Hygiene, *9*, 2, p. 466; March 1929.

PLATE I



A—

X-RAY PHOTOGRAPHS AT THE CLOSE OF EXPERIMENT

A. *Chick no. 1110 (Male, Pen 3) 90% cured, one and one third minutes daily exposure to a quartz mercury arc.*

UNIV. OF MICH.

PLATE I

B—

X-RAY PHOTOGRAPH AT THE CLOSE OF EXPERIMENT

B. *Chick no. 1148 (Male, Pen 2) no irradiation, condition—very severe rickets. These chickens both had severe rickets four weeks (28 days) earlier.*

UNIV. OF MICH.

MODES OF VIBRATION OF A ROUND PLATE CUT FROM A QUARTZ CRYSTAL

By A. M. Skellett

ABSTRACT

A large number of modes of vibration of a round quartz crystal plate of the "Curie" cut were photographed by means of the ionized gas appearing at the antinodal regions when the plate was mounted in an atmosphere of neon or argon at a few mm pressure and thrown into vibration piezo-electrically by a radio frequency oscillator.

Two general types of vibration of the plate were found. In the first kind the plate was symmetrically divided up into nodal and anti-nodal regions forming patterns peculiar to the exciting frequencies. In the second kind of which less evidence was found the vibration consisted of standing waves along chords or diameters of the plate with reflection at the edges. The crystal rested on a round brass plate which served as the lower electrode. Two different top electrodes were used, a point and a screen, the same pattern being photographed with each for the more prominent modes of vibration and from this it was concluded that the patterns were independent of the electrode arrangement. Evidence of vibration along axes oblique to the surfaces of the crystal plate was found. All vibrations observed were believed to be compressional in nature.

INTRODUCTION

It is well known that plates cut from natural quartz crystals with one of the major surfaces parallel to the optic axis of the quartz are active piezo-electrically. There are usually three[1] fundamentals of vibration, of which that through the thickness is generally used for frequency stabilization. For any particular plate however, Cady[2] has detected a large number of other frequencies at which vibration occurs by an audible method depending on the heterodyne note caused by the vibrating crystal and the slightly detuned exciting oscillator. A tabulation of these frequencies or spectrum of a given crystal has been made by Pierce.[3] That they correspond to as many different modes of mechanical vibration of the plate has been visually shown in an earlier paper[4] in which the method used herein is fully described.

EXPERIMENTAL ARRANGEMENT

The orientation of the electric axes of the quartz crystal in the plate is shown in Fig. 1a. The optic axis is perpendicular to the plane of these axes and parallel to the top surface. This is the "Curie" or "zero angle" cut, one of the electric axes being perpendicular to the surface of

[1] A. Hund, Proc. I.R.E., *14*, p. 452; Aug. 1926.
[2] W. G. Cady, J.O.S.A. & R.S.I., *10*, p. 480; 1925.
[3] G. W. Pierce, Proc. Am. Acad. Sci., *59*, p. 104; Oct. 1923.
[4] A. M. Skellett, J.O.S.A. & R.S.I., *17*, p. 308; Oct. 1928.

the plate. The crystal measured 2.80 mm thick by 25.10 mm in diameter. The faces were polished to transparency. The fundamental frequency of a vibration through the thickness of the crystal was 958.2 kilocycles corresponding to a wave length in air of 313 meters.

The crystal rested freely upon a round brass disc which served as the lower electrode with four small vertical glass tubes evenly spaced at the perifery to keep it in place. These tubes were spaced so that a movement of a millimeter or so in all directions was possible. Two top electrodes were used, a screen mesh and a point. The screen was used to approximate conditions that would obtain if another plate

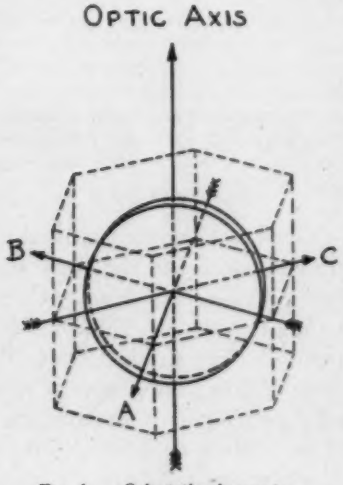


FIG. 1A. *Orientation in quartz.*

were used as a top electrode. It was made as flat and plane as possible and supported between the glass rods at about 0.5 mm from the surface of the crystal, a distance greater than the distance between the adjacent wires of the mesh. The relative position of the parts can be seen in Fig. 1B. A coil connected across the crystal electrodes was loosely coupled to the coil of a Hartley oscillator using a UX 210 tube with 400 volts on the plate. The coupling was variable so that the power supplied could be kept just below the value at which the gas in the tube became ionized by this exciting potential. A wave-meter which was calibrated with a crystal from the Bureau of Standards was very loosely coupled to the coils. Both neon and argon were used but most of the photographs were taken with argon. The mounted crystal was enclosed in a pyrex

tube with a plane glass window directly above the top surface. Pressures from 7 to 150 mm were used.

DISCUSSION OF FIGURES

When the plate was thrown into mechanical vibration by the impressed alternating voltage of the correct frequency for a particular mode of vibration, because of mechanical resonance and the piezoelectric effect, the antinodal regions of the surfaces became charged to alternating potentials higher than those applied so that it was at these regions that the gas became ionized. The configuration of the glow on the surface thus presented a beautiful picture of the manner in which the plate was vibrating. A large number of photographs were taken at various frequencies of the plate's spectrum and a selection of the best is shown in the accompanying figures.

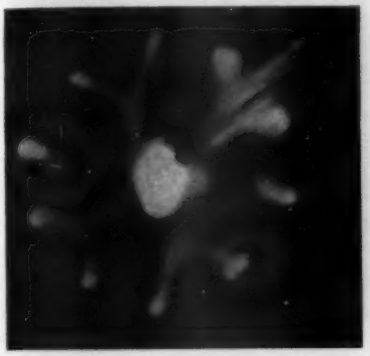


FIG. 1. $f=606$ kc.

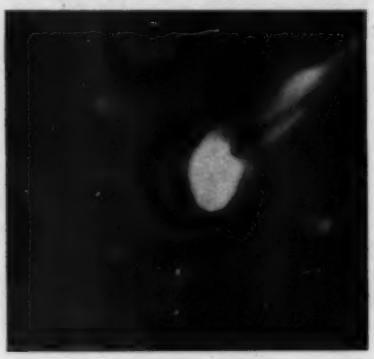


FIG. 2. $f=606$ kc.

Figs. 1 and 2 were taken but a few hundred cycles apart. In fact a movement of the hand towards or away from the coil of the oscillator would change the frequency by enough to bring in first one and then the other pattern. Now if these were the configurations of the antinodal regions which were 180° out of phase in the pattern of Fig. 3, then as the frequency was altered on either side of resonance they would shift phase with respect to the exciting potential and it is conceivable that this difference in phase might be great enough to cause one pattern to glow on one side of resonance and the other on the other side. The frequencies at which the glow occurred would not in general, however, be symmetrical about the resonant frequency. Attempts were made using a maximum of power and connecting the crystal directly across the

condenser of the oscillator to obtain the pattern shown in Fig. 3 but without success, probably because increasing the power very much caused a discharge throughout the whole tube. This photograph was obtained by exposing a plate first to the pattern of Fig. 1 and then slightly shifting the frequency to that of Fig. 2.

A similar phenomenon was noted on Fig. 4. When the frequency was shifted slightly another pattern appeared at the darker portions of the crystal with the regions which were light, dark. The frequency at which this pattern appeared as 727 kilocycles and if we take the velocity of sound in quartz perpendicular to the optic axis and divide it by this frequency we get a wave length in the plate of 7.40 mm. Measurement on the photograph of the distance between the extreme antinodal regions, reduced to the crystal dimensions and divided by 6,

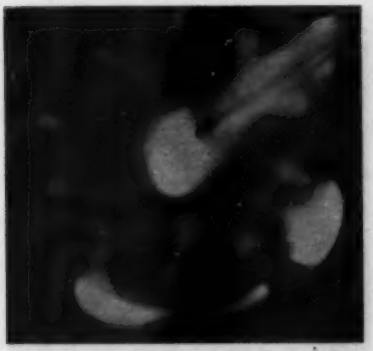


FIG. 3. $f=606kc.$

FIG. 4. $f=727$ kc.

the number of half wave lengths between them gave 3.70 which multiplied by 2 gave 7.40 mm in agreement with the calculated value so that this was evidently a set of standing waves across the plate. Where the distance across the standing waves between the edges of the plate was an even number of half wave lengths, the regions had a maximum brightness because of reflection at the edges. This was the type of vibration most prominent in the earlier investigation cited above. It would be expected that rectangular plates would give more of this type of pattern because of the possibility of more parallel antinodal regions of equal lengths spaced equidistant from one another.

It is difficult to determine which of the two types of vibration was present in Fig. 5. An average measurement across the antinodal regions gave a wave length in the crystal of 13.4 mm. The exciting frequency

was 388 kc which gives from the velocity of sound in quartz a wave length of 13.85 mm. The curvature of the lines of this pattern, however, would indicate a vibration of the type of Fig. 3 and the above agreement is not all that could be deisred.

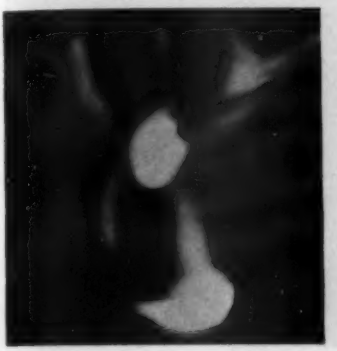


FIG. 5. *f=388 kc.*

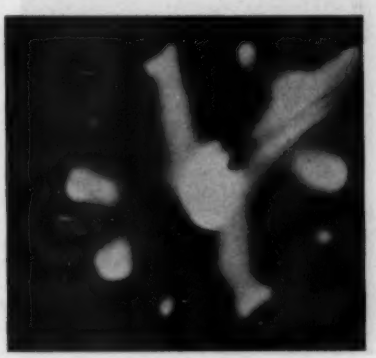


FIG. 6. *f=609 kc.*

Figs. 6 and 7 are probably patterns of the first type of vibration, i.e., of the same kind as shown by Fig. 3. Note particularly the similarity between Figs. 6 and 1 and the closeness of their frequencies.

FIG. 7. *f=427 kc.*

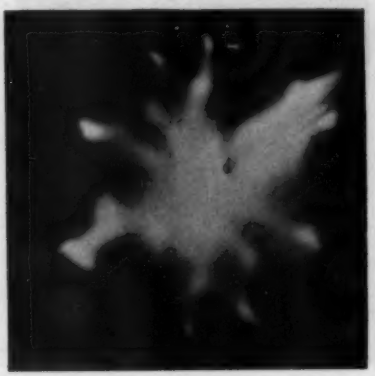


FIG. 8. *f=928 kc.*

Figs. 7, 8, 9 and 15 are representative of some of the other configurations photographed.

Fig. 10 was observed only on one occasion and photographed at that time. Contrary to the rest of the patterns it did not seem to be as sharp with frequency as did the others.

FIG. 9. $f=954$ kc.

FIG. 10. $f=574$ kc.

Figs. 11 and 12 are of the same pattern taken with the point and the screen electrodes. This was done to find out whether the shape of the electrode had any effect on the patterns. Other patterns were also ob-

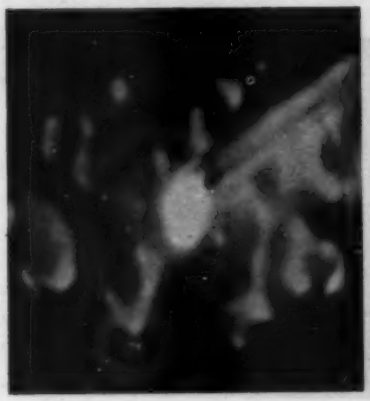


FIG. 11. $f=955$ kc.

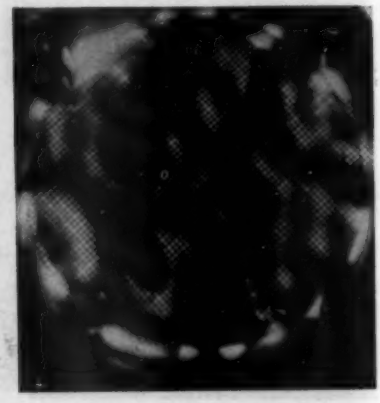


FIG. 12. $f=955$ kc.

served with both electrodes. From this it was concluded that the patterns were independent of the shape of the electrodes. It should be noted, however, that the ease with which any particular mode of vibra-

tion could be exicted would depend to some extent on the shape of the electrodes and their positions.

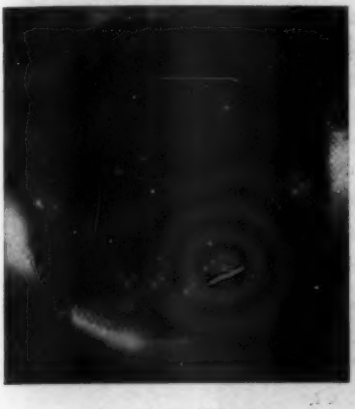


FIG. 13. $f=954.7$ kc.

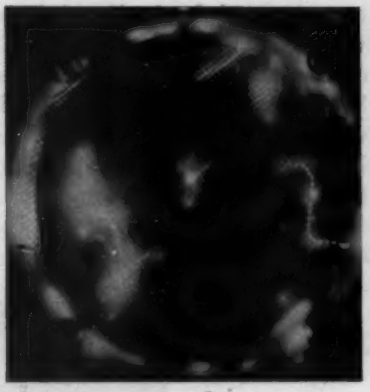


FIG. 14. $f=950$ kc.

This pattern was very easy to excite always appearing as its frequency was passed. The type of vibrations here evidently takes place along axes other than the three at right angles to one another.

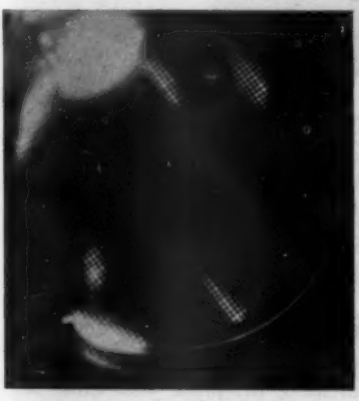


FIG. 15. $f=328$ kc.

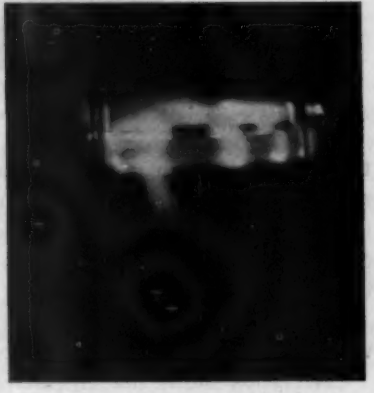


FIG. 16. $f=954$.

Evidence of this is seen in the edge-on view of Fig. 16 where the ionized gas is propelled away from the plate at an angle, in the positions of the spots on the edge of the crystal in Fig. 17, and in the small spot to the

lower right of the cross on Fig. 12 which was observed to be on the bottom of the crystal. If, as Fig. 16 seems to indicate, there is a vibration along an axis oblique to the surface of the crystal there would be a component of the motion of the surface parallel to the surface. This would shift the crystal sideways in its holder, an action which was observed many times. In fact by shifting frequencies the crystal could be made to hurl itself in either direction and at some frequencies it would turn. This shifting of vibrating crystals has been observed by Meissner[8] and others. Fig. 13 is representative of patterns similar to Fig. 12 of which there were a great number around the fundamental frequency of the crystal. An explanation of this large number is seen in that, were a mode of vibration at all possible, at the antinodal regions

FIG. 17. $f=955$.

there would be built up much greater amplitudes because of a double resonance effect, i.e., that due to the mode of vibration as shown in the figures and that due to the thickness being 1/2 of the wave length in the quartz so that there would be a node in the center (through the thickness) of the plate.

Hund[1] has observed a special case of this in which the mode of vibration was a harmonic of another fundamental.

Calculation shows that if different areas of the plate were different in thickness by as much as a wave length of sodium light their respective resonant frequencies would be different by as much as 200 cycles. The plate used here was polished to no such accuracy and it is probable that

[8] A. Meissner, Proc. I.R.E., *15*, p. 282; Apr. 1927.

it is because of this unevenness in the thickness that the figures at frequencies near the fundamental do not appear more regular. At other frequencies where double resonance does not occur the exactness of thickness is not such a critically important item and the effect is not so noticeable. This is borne out by Figs. 1, 2 and 6. Therefore, to accurately picture the modes of vibration, particularly around the fundamental, we probably will need a plate polished to an accuracy of better than 0.001 mm.

Figs. 13 and 14 show the appearance of patterns as the pressure of the gas was increased. It can be noticed that most of the glow was between the crystal and the top electrode because of the short mean free path of the gas. The positions of the bright lines on Fig. 14 probably cannot be taken as those of the antinodal regions since at higher pressures than a few mm there is a tendency for the centers of antinodal regions 180° out of phase to be connected with an arc-like discharge of this appearance. The pressure in this case was 8 mm of neon.

From the figures shown here it can be seen that most if not all of the patterns have a definite symmetry about an axis, which from the evidence of Fig. 4, is probably perpendicular to the optic axis. In other words it evidently lies in the intersection of the plane of the electric axes and the plane of the crystal plate.

SUMMARY

1. Two types of vibration of the quartz plate were found. In the first type the plate is symmetrically divided up in such a way that the antinodal regions form a definite pattern at the particular frequency at which this mode of vibration occurs. Adjacent antinodal regions are out of phase. At any instant if one antinodal region is in a state of compression, those on either side are in a state of rarefaction.

An understanding of this type of vibration can possibly be gained by considering Chladni's figures on metal plates, the fundamental difference of course being that these vibrations are compressional in nature while those observed by Chladni are flexural.

The other type of vibration is caused by standing compressional waves across the plate with reflection at the edges as observed with rectangular plates. It is probable that more modes of the first type of vibration would have been observed with the rectangular plate of the previous investigation had it been perfectly regular. Careful measure-

ment showed that at one end it was very slightly wider than at the other.

2. For the more prominent modes of vibration, the pattern of the antinodal regions of the vibrating plate was independent of the electrodes used and it is concluded that this holds for all modes of vibration.

3. Evidence was found to indicate that vibration may take place along axes oblique to the surfaces of the crystal plate.

UNIVERSITY OF FLORIDA,
GAINESVILLE, FLORIDA.

GRAPHIC SYMBOLS IN APPLIED OPTICS

By Noel Deisch

The accompanying draft of a system of conventions answering to the needs of instrumental optics is presented rather as a tentative or experimental formulation than as a guide for correct practice. Features of this list that may be noted, are (1) the device of using a given class character to distinguish certain of the larger families of symbols, and of progressively altering this parent character to secure the desired case differentiation, (2) the use of diacritical additions to express qualifications, such as material, structure, properties, orientation, or relationship, (3) the extensive inclusion of expressions denoting various types of mechanical freedom.

Experimental forms have naturally been introduced rather freely, and where a departure from established usage occurs it will be understood that the proposed character is included more out of respect for consistency than as an attempt at reform. All of these symbols, it is thought, meet the requirements of suggestiveness, saliency, and symmetry, of being easily drawn, and of being reproducible at great reduction by the etched zinc process of line engraving.

In the following description of the figures, (*Ill.*) denotes that the symbol in question is an illustrative modification of the preceding type symbol: it may of course be only one of several possible modifications. Material within brackets refers to the derivation of the symbol, and is included to promote the forming of mnemonic associations.

SOURCES

1. Source. Generalized symbol,—character of source not defined.
2. Point source.
3. (Ill.) Incandescent tungsten bead (e.g., Philipps lamp.) The lines representing wire terminals in this and in some of the following illustrations show how the symbol is to be included in circuit diagrams.
4. Line source.
5. (Ill.) Carbon filament lamp.
6. Broad-line source.
7. (Ill.) Ribbon filament lamp.
8. Arc. [Tip of carbon arc electrodes.]
9. (Ill.) Iron arc.

10. Spark source. [Spherical discharge electrodes.]
11. (Ill.) Liquid-quenched zinc electrode spark source. [Liquid is indicated by short horizontal lines, following conventional practice.]
12. Gas discharge tube.
13. (Ill.) Neon gas discharge tube.
14. Metallic vapor lamp. [Schematic mercury vapor lamp.]
15. (Ill.) Cadmium amalgam metallic vapor lamp.
16. Combustion source. [Fan-shaped flame.]

RECEPTORS

17. Generalized symbol,—character of response not stated.
18. Photo-electric cell.
19. (Ill.) Potassium cathode cell.
20. Glow-discharge tube, used as receptor.
21. Photovoltaic cell,—electrolyte cell. [Electrode dipping into liquid.]
22. (Ill.) Sodium iodide photovoltaic cell.
23. Cell employing photosensitive semiconductor, (e.g., selenium cell.)
24. (Ill.) Thallium sulphide, (e.g., "Thallofide") cell.
25. Bolometer. [Resistance wire of bolometer.]
26. Thermocouple. [Soldered alloy wires.]
27. Radiometer. [Diamond-shaped vanes of demonstration radiometer.]
28. Eye.
29. Photosensitive surface,—generalized symbol. [Hurter and Driffield cipher.]
30. Photographic plate. [Silver halide.]
31. Fluorescent screen. (The same symbol may of course be used when the fluorescent screen is employed as a source.) [Wavy line symbolic of change of wave-length.]
32. (Ill.) Calcium tungstate fluorescent screen.

LENSES

Lenses are rather difficult to draw, and hence unless the specification of the construction or of the specific function of the lens is of importance, the forms showing simple converging or diverging lenses will be used, even if the combination in question actually be of high complexity. Achromatism is shown by a dividing line. By way of example a photo-

graphic lens in association with its iris and shutter is indicated at 37, whereas at 38 the construction of an anastigmat is detailed. So also at 39 the coddington type of simple magnifying lens is shown, whereas ordinarily this would be drawn as indicated at 33. A description of the mechanical symbols forming part of the characters 37 and 38 and of some other characters in the list will be found under the discussion of symbols of mechanical freedom.

33. Converging lens; (A) uncorrected, (B) achromatic.
34. Diverging lens; (A) uncorrected, (B) achromatic.
35. Cylindrical lens, uncorrected. (A) Axis normal to plane of drawing, (B) axis parallel to plane of drawing.
36. Cylindrical lens, achromatic. (A) Axis normal to plane of drawing, (B) axis parallel to plane of drawing.
37. (Ill.) Achromatic lens in combination with iris and shutter.
38. Anastigmat lens in combination with iris diaphragm and shutter.
39. Coddington lens.

PRISMS

40. Polarizing prism.
41. Nicol prism.
42. Crossed polarizing prisms.
43. Polarizing prism with azimuth of plane of polarization specified. (Crossed prisms may of course also be indicated by this means, i.e., 0°, 90°.)
44. Prism with transmitting-reflecting interface.
45. (Ill.) Prism with half-platinized interface.
46. Photometric prism.
47. Lummer-Brodhun photometric prism.

MIRRORS

48. Speculum,—solid metallic mirror.
49. Full surface metallized mirror.
50. (Ill.) Full surface platinized mirror.
51. Various forms of transmitting-reflecting mirrors. (A) Generalized symbol, without specification of material or transmission-reflection characteristics, (B) half-silvered mirror, (C) partly silvered mirror transmitting 60% of incident light.
52. Banded or barred mirror.
53. Diffraction grating. [Lines of grating indicated.]
54. Echelette grating.
55. Quarter-wave plate.

APERTURES

56. Aperture, type unspecified.
57. (Ill.) Adjustable aperture, type of aperture and method of adjustment unspecified.
58. (Ill.) Adjustable aperture, slidable closure.
59. Adjustable radial diaphragm, scissors closure.
60. Adjustable radial diaphragm, sector closure.
61. Zonal diaphragm.
62. Variable zonal diaphragm, (Iris.)
63. Artificial pupil.
64. Rotating sector, fixed opening.
65. Rotating sector, adjustable opening.

SLITS

66. Slit; (A) axis of slit normal to printed sheet, (B) axis of slit parallel to printed sheet.
67. (Ill.) Slit, opening of fixed width.
68. (Ill.) Slit, opening of adjustable width.
69. (Ill.) Slit, opening slidably adjustable, and means provided for making a reading of the setting.

SCREENS

70. Diffusing screen, either transmitting or reflecting; (A) thin, (single row of dots), (B) thick. [The dots are suggestive of diffusing colloidal particles.]
71. Selective material, or color filter. [Hatching of various kinds is used to indicate colors in heraldry and in Patent Office practice; perhaps different types of hatching indicating different colors might be useful in certain optical diagrams, such as those involving the apparatus of color photography.]
72. (Ill.) Filters in which the transmission characteristics are suggested. (A) Filter excluding ultra-violet, (B) filter having maximum transmission at $\lambda = 480$, (C) monochromat filter transmitting the line 546 (of the mercury spectrum.) An arrow directed toward the source indicates excluded radiation, directed away from the source it indicates transmitted radiation. For example, in 72 (C) above, if the filter were a rejecting monochromat, such as the interference filter described by Herbert E. Ives,[1] the fact would be shown by directing the arrow against the source.

[1] J.O.S.A., *1*, pp. 49–63; 1917.

73. Liquid filter and filter-cell.
74. Neutral density.
75. Neutral wedge.
76. Screen density,—wire or half-tone screen. [This is the symbol employed by the Patent Office to designate a wire screen.]

VARIOUS

77. Reticle, (single wire.)
78. Reticle, (crossed wires.)
79. Filar micrometer.
80. Graticule,—lines engraved or etched on glass.
81. Electro-optic cell,—Kerr cell.
82. Magneto-optic cell,—Faraday cell.
83. Shutter.

SYMBOLS OF MECHANICAL FREEDOM

The considerable inclusion of mechanical elements in optical instrument diagrams demands that attention be given in this connection to the expression of various mechanical relationships and movements, and since no extensive or systematic graphic symbology appears to have been developed in connection with the science of kinematics, it has been necessary to use a list which cannot claim the support of any precedent, and hence may quite possibly give an impression of being somewhat heteroclyte.

The method here suggested of specifying the six degrees of mechanical freedom is based on the use of four characters, intended to be attached to the parent symbol as subscripts, and in each case referring specifically to a given one of the three axes of rotation or translation, but expressing freedom in the most general way and without local constructional connotation. These characters are shown in column 84 of the section headed *Symbols of Mechanical Freedom* in the drawings. The upper half of this column, as indicated in the legend, is concerned with simple rotational freedom, and the lower half with simple linear freedom. The character referring to angular freedom, as used in cases where the axis of rotation lies in the plane of the printed sheet or more particularly on the x or the y axis, consists of a cross with one thick and one thin arm, suggestive of a wheel and arbor in side elevation, the arbor, corresponding to the narrower line, being of course drawn parallel to the axis about which rotation in the member which it modifies takes place. The character expressing angular freedom about an axis lying normal to the plane

of the paper, i.e., on the z axis, may be conceived as an end elevation of the above symbol, that is, it represents in the simplest manner a wheel and arbor in end elevation.

The symbols in the lower half of column 84, which express linear freedom, suggest a slide-bar or guide depicted in simplest outline. The first two characters have reference to motion along axes in or parallel to the plane of the sheet, and the third, which may be taken as an end elevation of the former, to motion along an axis normal to the plane of the sheet.

Where it is desired to show that the adjustment is readable and controllable quantitatively by means of a graduated scale, the fact is expressed in any of the above cases by combining with the symbol of freedom a short mark representing an index or fiducial line, as shown in column 85 of the table.

SIMPLE FREEDOM

84. Parent symbol indicating simple mechanical freedom.
85. (Ill.) Same as 84, but with addition of a modifying mark showing that a graduated scale is provided by which a reading may be taken of the setting.

MODIFIED FREEDOM

86. The permitted motion or range of adjustability is restricted between two limits.
87. The motion (usually abrupt) is from one to another of two limits, or positions of repose, (e.g., blades of a photographic shutter.)
88. The member may be moved from one to another of two locked positions. (A plurality of locked positions may of course be shown if required.)
89. Freedom extends indefinitely, or is produced beyond limits of drawing.
90. Universal joint.
91. Three-way slide.
92. Complete freedom.
93. Fixity.

In the occasional case where it is necessary to indicate the position in the member modified of the axis about which rotation occurs, this may be done by means of a center line, or, where the axis is normal to the plane of the paper, by intersecting center lines, as shown in the lower figure of column 96.

It is sometimes convenient to show adjustability as disassociated from the mechanical context, and for this purpose the oblique arrow, which has been applied in this sense rather widely in electrical circuit diagrams, has been introduced here. However, since the shaft of the arrow would cut the optic axis at an inconvenient point if drawn across the symbol in the manner usual in electrical diagram practice (this would be especially true, for instance, in the case of slits or other narrow apertures) its use has been made uniform with that of the other modifying symbols by removing it to the position of a subscript. Examples of the application of this symbol are shown at 57, 62, and elsewhere in the drawings.

When the permitted movement is provided to allow the permanent or semipermanent adjustment or calibration of the instrument, but is not employed during its regular use, the modifying symbol is enclosed in parentheses. An example is shown at 103, where the nicol prism 9 is shown as normally set at an angle of about 45° with relation to the prisms 12 and 13, but may be moved slightly to one side or the other of this position to secure a proper balance of the fields, as indicated by the oblique arrow in parentheses.

The more specialized symbols expressing modified freedom will of course be only occasionally in demand, and will be brought into requisition only where the meaning would not otherwise be sufficiently specific.

A comparison of these symbols with conventional usage is shown under *Alternative Forms of Statement*. The examples given in the upper row of figures represent typical optical diagram practice, as selected from the literature, whereas the symbols in the lower row represent the same elements expressed in the proposed system. Though it is quite evident that all of the symbols shown in the upper row have the merit of being readily interpretable, on the other hand certain of these are difficult to draw, and are so specific as to represent only a case solution. Some morevoer appear to be out of character with the best diagram practice by being rather too pictographic.

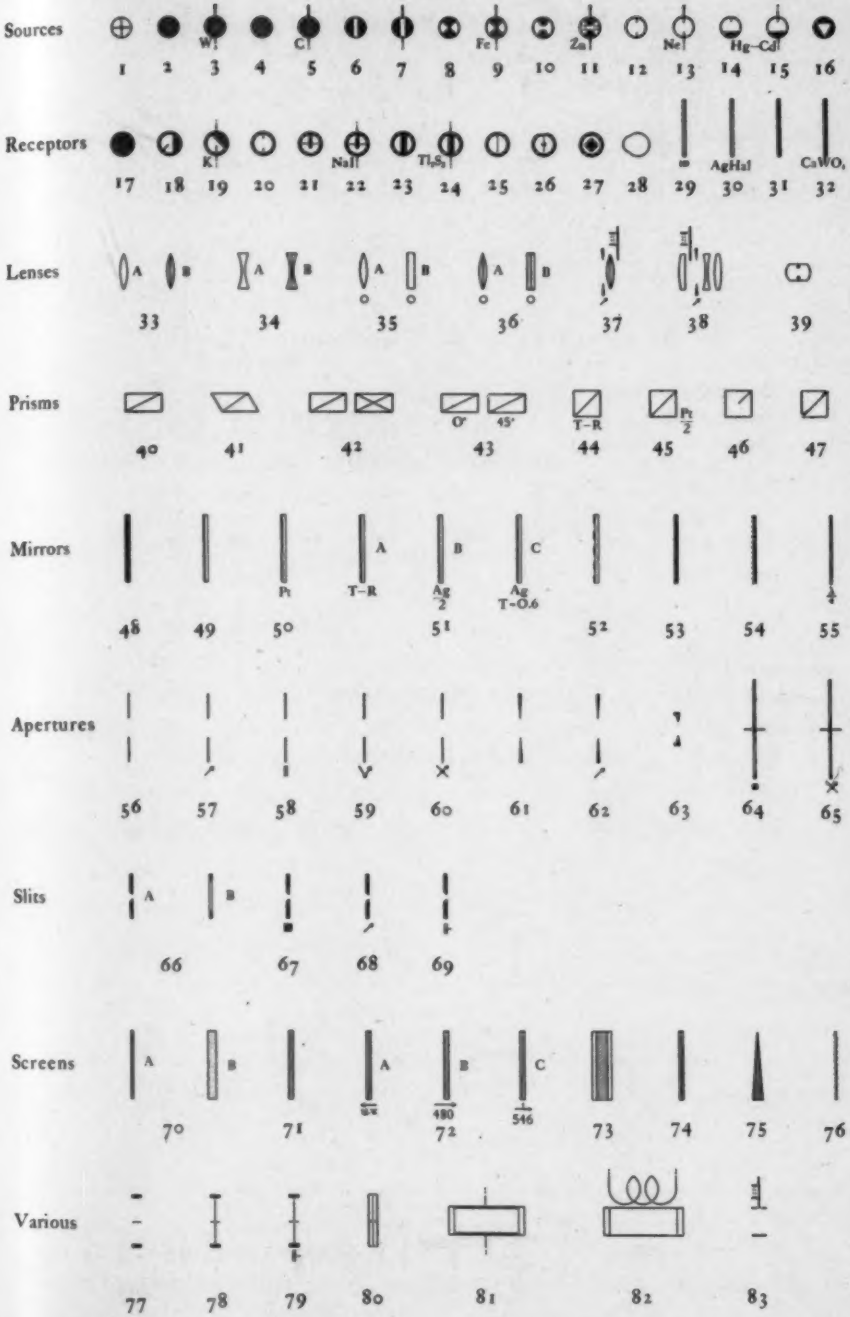
ALTERNATIVE FORMS OF STATEMENT

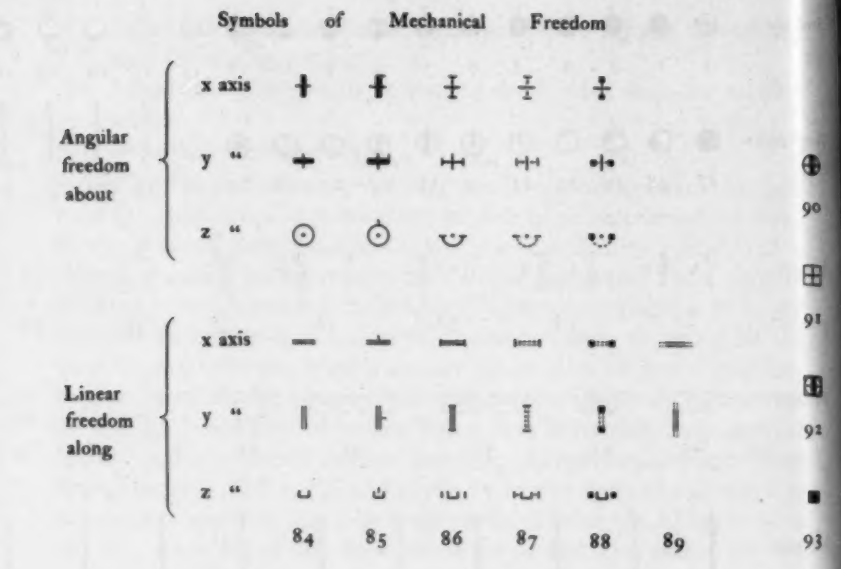
94. Adjustable slit.
95. Filar micrometer.
96. Spectrometer prism with reading drum.
97. Angularly adjustable prism, with scale for reading setting.

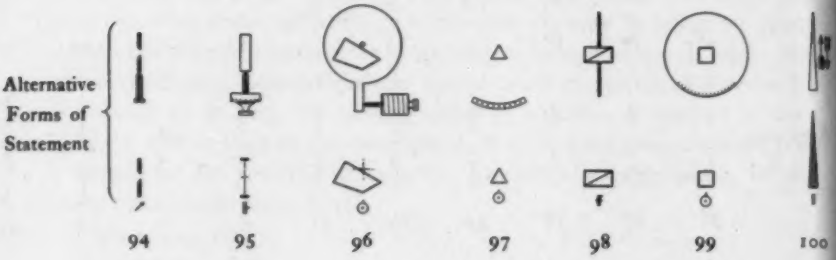
98. Adjustable polarizing prism,—axis parallel to printed sheet, with scale for reading setting.
99. Adjustable prism,—axis normal to sheet, with scale for reading setting.
100. Adjustable neutral wedge.

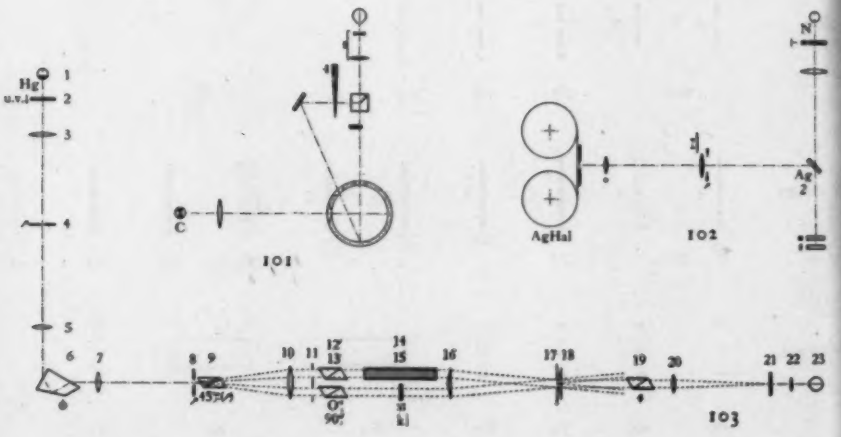
At 101, 102, and 103 are given diagrams of optical apparatus to illustrate the application, in writing the formula of an instrument, of some of the symbols here presented;—101 shows a density meter, 102 a recording optical extensometer, and 103 an apparatus (here slightly modified to accommodate it to the page) designed long ago by Professor Straubel for measuring the ultra-violet transmissivity of Jena glass. In these figures, following what is already the practice in certain quarters, the line of collimation or axis of the optical system has been represented by the same character of broken line as that universally employed for indicating center lines in machine drawing. Reference characters (used illustratively in 103) are placed above or to the right of the main or substantive symbols, the adjectival particles being placed below or to the left, according as the direction of the optical axis makes the one or the other placement the more convenient. The parts are numbered in progressive order from source to receptor. It may be remarked that not a little time and annoyance in locating reference numerals might be saved the reader, especially where complicated apparatus is described, by seeing to it that the natural order of numbers is applied in the diagram rather than in the description, that is, that the numerals are arranged on the drawing in columns, in clockwise progression, or in some other consecutive form.

WASHINGTON, D. C.,
JANUARY 21, 1930.


Sources
W
C
Fe
Zn
Ne
Hg—Cd
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
Receptors
K
NaI
Tl₂S
AgHal
CaWO₄
17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32
Lenses
A
B
33 34 35 36 37 38 39
Prisms
0°
45°
T–R
Pt
40 41 42 43 44 45 46 47
Mirrors
Pt
T–R
Ag
T–0.6
A
B
C
48 49 50 51 52 53 54 55
Apertures
56 57 58 59 60 61 62 63 64 65
Slits
A
B
66 67 68 69
Screens
A
B
C
546
480
70 71 72 73 74 75 76
Various
77 78 79 80 81 82 83


Symbols of Mechanical Freedom
Angular freedom about
x axis
y "
z "
Linear freedom along
x axis
y "
z "
84
85
86
87
88
89
90
91
92
93


Alternative Forms of Statement
94
95
96
97
98
99
100


Hg
u.v.
C
AgHal
Ag
101
102
103